JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1993, 59, 29-60 NUMBER 1 (JANUARY)

DETERMINANTS OF HUMAN PERFORMANCE ON CONCURRENT SCHEDULES

P. J. HORNE AND C. F. LOWE

UNIVERSITY COLLEGE OF NORTH WALES

Six experiments, each with 5 human adults, were conducted to investigate the determinants of human performance on multiple concurrent variable-interval schedules. A two-key procedure was employed in which subjects' key presses produced points exchangeable for money. Variables manipulated across experiments were (a) changeover delay (Experiments 2, 4, and 6), (b) ordinal cues related to scheduled reinforcement frequencies (Experiments 3 and 4), and (c) instructions describing the ordinal relations between schedule-correlated stimuli and scheduled reinforcement frequency (Experiments 5 and 6). The performances of only 13 of the 30 subjects could be described by the generalized matching equation and were within a range of values typical of those reported in the animal literature. Eight subjects showed indifference, 9 undermatched, 7 approximated matching, 3 overmatched, and a further 3 responded exclusively to the richer component of the concurrent schedules. These differing modes of responding were closely related to the different types of performance rules reported by subjects in postexperimental questionnaires. The results are in good agreement with those from studies of human performance on single schedules, suggesting that rule-governed behavior, in interaction with contingencies, may be an important determinant of human choice.

Key words: matching law, choice, multiple concurrent variable-interval schedule, instruction, schedule-correlated stimuli, rule-governed behavior, verbal reports, key pressing, adult humans

---

There have been numerous reports in the literature of marked differences between the performances of adult humans and other animal species on schedules of reinforcement. This is true on fixed-interval (FI) and fixed-ratio (FR) schedules, and is evident both in response patterning and in sensitivity to the schedule parameters. The evidence also suggests that the occurrence of rule-governed and other verbal behavior in humans may give rise to some of these differences. (For reviews of the literature, see Lowe, 1979, 1983; Matthews, Shimoff, Catania, & Sagvolden, 1977.)

A series of studies (Bentall & Lowe, 1987; Bentall, Lowe, & Beasty, 1983, 1985; Lowe, Beasty, & Bentall, 1983) has demonstrated that the operant behavior of preverbal infants performing on FI, FR, and differential-reinforcement-of-low-rate (DRL) schedules has few of the distinctive features of human adult performance but is instead indistinguishable from that of nonhumans. By the time children reach the age of 2 to 4 years, however, and begin to emit some verbal behavior, their performance on FI schedules is no longer like that of nonhumans, but neither does it resemble the behavior of adults. At 5 or 6 years of age, when they have acquired verbal skills that enable them to describe schedule contingencies and to formulate rules for responding, they behave much as adults do, with similar response patterning and similar insensitivity to alterations in schedule of reinforcement. That the acquisition of verbal behavior, including rules, is the variable responsible for these age-related changes in schedule performance is not demonstrated unequivocally by these experiments; other explanations are always possible. But that such is indeed the case is supported by the finding that the developmental sequence can be accelerated by appropriate verbal instruction (see Bentall & Lowe, 1987; also Bem, 1967; Luria, 1961). This also accords with the results of several studies that have demonstrated the effects both of verbal instructions and of what subjects say on what they do (see Catania, Matthews, & Shimoff, 1982; LeFrancois, Chase, & Joyce, 1988; Michael & Bernstein, 1991; Paniagua & Baer, 1982).

Some of these findings were reported at the First European Meeting on the Experimental Analysis of Behaviour held at the University of Liege, Belgium, in 1983. During the study P. J. Horne was in receipt of a graduate studentship from the University College of North Wales, Bangor. Thanks are due to Llio Ellis Williams and Patricia Lowe for their help in preparing this manuscript. Send correspondence and reprint requests to Pauline Horne or Fergus Lowe, Department of Psychology, University of Wales, Bangor, Gwynedd LL57 2DG, United Kingdom.

Although much of the literature has consistently shown a number of differences between the performance of nonhumans and verbally accomplished humans, there is one area of human operant research that has yielded more equivocal data. This work has centered around Herrnstein's (1970) equations, particularly the matching law and its derivations, which aim to quantify the effects of reinforcement on behavior. A number of studies have claimed that human performance on variable-interval (VI) schedules, whether presented singly or concurrently, is characterized, just as nonhuman performance is, by a lawful relationship between rate of responding and rate of reinforcement that closely conforms to these equations (e.g., Bradshaw, Ruddle & Szabadi, 1981; Bradshaw, Szabadi, & Bevan, 1976, 1977, 1979; Bradshaw, Szabadi, Bevan, & Ruddle, 1979; Ruddle, Bradshaw, Szabadi, & Bevan, 1979). Indeed, some authors have been so impressed by this evidence that they have suggested that contemporary research with nonhumans in this area may be of direct relevance to applied behavior analysis, with profound implications for the study of human economic behavior and the behavior of people in clinical, educational, and social settings (see Epling & Pierce, 1983, 1988; Hamblin, 1979; McDowell, 1981, 1982; Pierce & Epling, 1980, 1983; Rachlin, 1980; Schwartz, 1984; Winkler, 1980). Certainly these findings, or at least the common interpretation of them, seem at odds with much of the other human operant literature, and they are potentially problematic for those accounts that suggest that variables such as rule-governed behavior operate at the human but not at the nonhuman level.

Within the context of animal choice behavior, Herrnstein (1970) proposed that when two explicitly programmed sources of reinforcement, A and B, are available independently and concurrently, the allocation of responding to these two sources will be described by the following hyperbolic equations:

$$R_A = \frac{Kr_A}{r_A + r_B + r_0} \tag{1}$$

and

$$R_B = \frac{Kr_B}{r_A + r_B + r_0}, \tag{2}$$

where R_A and R_B are the number of responses per session to Alternatives A and B, respectively, and r_A and r_B are the obtained frequencies of reinforcement for the alternative responses. K and r_0 are empirically derived parameters, where K is response rate at asymptote; r_0 is the rate of reinforcement at half-maximal response rate and is theoretically equivalent to all implicit sources of reinforcement in the experimental setting (Herrnstein, 1970).

When K and r_0 are invariant, Equations 1 and 2 may be combined to derive the matching law, which states that relative response rates approximately equal the relative reinforcement frequencies for the two alternatives. This is described mathematically as follows:

$$\frac{R_A}{R_A + R_B} = \frac{r_A}{r_A + r_B}. \tag{3}$$

Several studies of animal performance on concurrent schedules have reported deviations from the ideal matching relationship (cf. Baum, 1979; de Villiers, 1977) and, following Staddon's (1968) earlier suggestion that choice behavior might be best described by a power function, Baum (1974) proposed a modification to Equation 3 that has become known as the generalized matching law:

$$\frac{R_A}{R_B} = k\left(\frac{r_A}{r_B}\right)^a, \tag{4}$$

where R_A, R_B, r_A, r_B are as defined in Equations 1 and 2, and k and a are two free parameters, estimated empirically, usually by fitting a straight line to a logarithmic transformation of response and reinforcement ratios. The parameter a provides an index of the sensitivity with which response ratios change when reinforcement ratios are varied. If the value of a is greater than 1, as a result of a disproportionately greater preference for the richer reinforcement schedules, performance is described as overmatching. In the case of a disproportionate preference for the poorer schedules, a is less than 1 and there is said to be undermatching. The parameter k provides an index of bias, and a value of k different from 1 signifies a constant proportional bias toward one alternative as opposed to the other (e.g., a positional bias). When both a and k are equal to 1, "ideal" or "perfect" matching is said to obtain and Equation 4 reduces to Equation 3.

Many studies of concurrent schedule performance have employed a changeover delay (COD); this is a time period during which no scheduled reinforcement is delivered for a specified time following a change from responding on one alternative to responding on the other, and it is designed to prevent reinforcement of changeover responses. A number of authors have argued that the inclusion of a COD is an important feature of any effective concurrent schedule procedure (see Catania, 1966; McSweeney, Melville, Buck, & Whipple, 1983).

Several reviews of the literature have shown that performance of nonhumans conforms closely to the generalized matching law, with values of a falling between 0.5 and 1.3 in the great majority of cases (Baum, 1979, 1983; de Villiers, 1977; Horne, 1986; Wearden & Burgess, 1982), though procedural variables have been found to affect sensitivity (Taylor & Davison, 1983; Todorov, Oliveira Castro, Hanna, Bittencourt de Sa, & Barreto, 1983). Even when departure from ideal matching occurs, however, the generalized matching law almost invariably accounts for a high proportion of the data variance (see de Villiers, 1977). A number of instances of matching have been reported in human studies (e.g., Baum, 1975; Buskist & Miller, 1981; Conger & Killeen, 1974; Ruddle et al., 1979; Schroeder & Holland, 1969) but also some gross departures (e.g., Navarick & Chellsen, 1983; Oscar-Berman, Heyman, Bonner, & Ryder, 1980; Pierce, Epling, & Greer, 1981; Schmitt, 1974; Takahashi & Iwamoto, 1986; Wurster & Griffiths, 1979). There are many procedural differences among studies of human choice, however, that make it difficult to account for the variability in the results both within and between experiments.

The most widely cited research in the area of human choice has been conducted by Bradshaw and colleagues who, in a number of studies, have reported conformity of human choice performance to the generalized matching law comparable to that found in nonhumans (Bradshaw et al., 1976, 1977, 1981; Bradshaw, Szabadi, & Bevan, 1979; Bradshaw, Szabadi, Bevan, & Ruddle, 1979). Their studies employed a five-ply multiple concurrent VI VI procedure wherein each of five different VI pairs were alternated with 5-min rest periods. In one version of the procedure (Bradshaw et al., 1981; Ruddle et al., 1979), successive 10-min presentations of each of the following VI schedules occurred in pseudo-random order on the right-hand mandipulandum (Component A): 8 s, 17 s, 51 s, 171 s, and 720 s. A VI 51-s schedule was concurrently operative on the left-hand manipulandum (Component B). Each VI schedule on the right-hand manipulandum was signaled by a different stimulus in the form of one light among a row of five being illuminated for the 10 min that the schedule was in operation. An alternative version of the Bradshaw procedure, as outlined in, for example, Bradshaw et al. (1976), entailed the use of a single response manipulandum and a changeover button that enabled subjects to switch between components. No COD was employed in any of the experiments. The subjects were human adults, and responding was reinforced during all 16 sessions by the delivery to a counter of points that were later exchanged for money.

An interesting procedural aspect of all these studies concerns the stimuli that were used to signal the schedules in operation. The position of the light illuminated in the row of five lights was *ordinally* related to the reinforcement frequency signaled. Thus, when the VI 8-s schedule was operative, the light furthest to the left was lit; when VI 17 s was operative, the light second furthest to the left was lit, and so on. Moreover, for the first experimental session, the VI schedules were presented only on the right-hand manipulandum (or Component A), the left-hand manipulandum (or Component B) being inoperative for that session, and these schedules were presented in an ascending order of interval value. It is possible that this ordinal information, albeit implicit, may have led subjects to describe the schedules and formulate rules for responding in such a way as to favor the emergence of matching-like behavior. For example, the schedule signaled by the left-most light (VI 8 s) might have been described as "best" and those to the right as progressively "worse." Because these studies have produced some of the most convincing evidence for matching in human performance and show an apparent conformity of human and nonhuman performance that appears to be at odds with much of the rest of the human operant literature, we decided to carry out a series of experiments that attempted to replicate the Bradshaw findings and, in addition, to investigate the factors responsible for human choice under these conditions. For these reasons our experimental procedure follows that employed in the Bradshaw studies, except where variables of particular interest are systematically manipulated.

GENERAL METHOD

The general subject, apparatus, and procedure specifications apply to all experiments unless exceptions are noted.

Subjects

Thirty normal adults participated in six experiments, 5 subjects being assigned to each experiment. They were between 17 and 35 years of age and were from various occupational backgrounds; they had no previous experience with psychology experiments. Subjects were recruited by means of advertisements in the local area and through personal contacts. The initials identifying subjects do not correspond to their real names.

Apparatus

The experiments were conducted in a cubicle with a floor area of 240 cm by 189 cm. The response console, which was mounted on one wall, contained two Lehigh Valley human response panels placed 24 cm apart. A points counter was mounted midway between the response panels, which were operated by a force of 6 N, producing an audible "click." The experiments were controlled by, and data were recorded on, an Apple® II microcomputer situated outside the cubicle.

Procedure

In the first session only the right-hand panel, hereafter referred to as Key A, was operative. The following instructions (cf. Bradshaw, Szabadi, & Bevan, 1979) were given to the subjects in Experiments 1, 2, 5 and 6:

> This is a situation in which you can earn money. You earn money simply by pressing the right-hand key. Sometimes when you press the key the counter will click and increment one point. This means you will have earned one penny. The total amount you have earned is shown on the counter. You can tell when you have pressed the key properly by listening for a slight click coming from inside the key. When the key is lit up with a shape, it means that you are able to earn money. At the beginning of the session, the key will be lit for 10 minutes: throughout this time you can earn money. At the end of 10 minutes the key will go blank and will remain like this for 5 minutes. During this time you may rest. After the rest period the key will be lit up with another shape for 10 minutes, and so on, until six different shapes have been presented. The order of the shapes with which the key is lit will vary from day to day. At the end of each session we will take the reading from the counter and note down how much you have earned. You will be paid in a lump sum at the end of the experiment. If you are right-handed use only your right hand for key pressing; if you are left-handed use only your left hand for key pressing.

In Experiments 3 and 4 these and subsequent instructions were altered by replacing the references to "shapes" with "row of dots."

Six different constant-probability VI schedules (Catania & Reynolds, 1968), each consisting of 30 intervals in a randomized sequence, were in operation on Key A. As in the Bradshaw procedure, no further reinforcers could be scheduled until a reinforcer that had been set up had been collected. During the operation of each VI schedule, which lasted 10 min, the response panel was illuminated with the appropriate stimulus; in four of the experiments (1, 2, 5 and 6) these stimuli were geometric shapes. The following schedules and their associated stimulus shapes were employed: VI 10 s (+); VI 20 s (○); VI 50 s (△); VI 175 s (—); VI 500 s (⊙); VI 720 s (□). (The stimuli used in Experiments 3 and 4 are described in the Method section of those experiments.) The 10-min schedule presentations alternated with 5-min rest periods. The order of presentation of VI schedules in the first session was VI 50 s, VI 10 s, VI 20 s, VI 175 s, VI 720 s, and VI 500 s.

In the second session, the left-hand panel (Key B) was also operative, and the following instructions were given to the subjects:

> As you can see there is a key to the left of the counter. This key has one shape associated with it. While the key is lit up with this shape you have an opportunity to earn money by pressing the key. As with the other key, sometimes when you press, the counter will click over and increment one point: this means you will have earned one penny. This key will be lit up throughout the session apart from rest periods.

> If at any stage, while pressing the right-hand key you think it might be a good idea to change to this key, you may do so and you may switch back again any time you wish. Thus throughout the session you have a choice between two keys. If you are right-handed, use only your right hand for key pressing; if you are left-handed, use only your left hand for key pressing.

Thus, in addition to the six schedules in operation on Key A, a VI 50-s schedule was concurrently operative on Key B, which was illuminated by the same stimulus that signaled VI 50 s on Key A (e.g., the triangle, when geometric shapes were used). In any one session, the six concurrent schedules were presented in a quasi-random sequence, with the constraint that each occurred in a different ordinal position on successive days. In line with the Bradshaw studies, no COD was employed in three of the six experiments (i.e., Experiments 1, 3, and 5); in the remaining experiments a 3-s COD was operative, and responses emitted during the first 3 s following a changeover response had no scheduled consequences.

Each reinforcement consisted of the addition of 1 point to the counter; the monetary exchange rate was the same as that used in the Bradshaw studies (i.e., 1 penny per point earned). This yielded average earnings per session of approximately £1.50 ($2.55) to £1.80 ($3.06). The instructions to subjects described above were closely modeled on those of Bradshaw, Szabadi, and Bevan (1979), with the addition that all subjects were asked not to speak to anyone other than the experimenter about the study. As in the Bradshaw procedures, 16 1.5-hr sessions were conducted with each of the subjects.

At the end of the experiment, subjects were asked to complete a questionnaire consisting of 15 main questions, eight of which branched to one or two additional questions in the event of a "yes" response. The questionnaire was designed to present subjects with the opportunity to describe the scheduled contingencies and to give an account of the factors that they considered determined their responding. In order to minimize biasing of subjects' verbalizations, the questionnaire was organized in such a way as to present questions of a very general nature initially (e.g., Q.1, "What did you think the experiment was about?"). Subsequent questions were progressively more focused (e.g., Q.3, "What was your reason for continuing to press on the keys?" and Q.4, "How did you set about winning points?"). Subjects were then asked in Q.6 whether, and if so why, they had any preferences among the schedule stimuli; in Q.8 they were asked whether they considered that the schedule stimuli had any particular relationship to the points they earned, and if so, what they thought that relationship was.

Following questions aimed at investigating the possibility of a general position bias in response allocation, more detailed questions (Q.11 through 13) were presented, inviting subjects to describe how they distributed their responding across the operanda during a session and, if this changed in the course of the experiment, to provide a chronological account of their responding. Finally, subjects were asked whether they had any other information that might have a bearing on how their performances might be understood.

Upon completion of the six experiments, subjects' postexperimental questionnaire responses were analyzed with a view to (a) identifying and categorizing any performance rules they might have used to guide their behavior and (b) identifying their order of preference for the schedules. First we delineated five categories of choice performance, each defined in terms of a specified range of generalized matching equation slope values as follows: *indifference* (−0.20 to 0.25); *undermatching* (0.30 to 0.75); *approximate matching* (0.80 to 1.25); *overmatching* (>1.3); and *exclusive preference* (responding allocated to the richer alternative on all six concurrent schedules). Subjects' verbal reports were then sorted "blind" by one of the experimenters (P. J. Horne), each subject's performance rules being assigned to one of five corresponding categories on the basis of the performance that should ensue were the rules to be followed. For example, a subject might report that she followed a performance rule to the effect that the higher the schedule value, the more she should respond. On the face of it, were this rule to be followed, a performance close to matching should result, so the rule would be classified as approximate matching. On the other hand, a subject might report that because she had no preference between schedules, she decided to allocate responding randomly across the keys; this would be classified as an indifference rule. Not all rule statements were as straightforward as these, however (see Table 1 for examples in each category), and

Table 1

Examples of performance rules assigned to each of the five categories.

Category	Examples of performance rules
Indifference	Subject TM, Experiment 1: I tried several strategies including pressing particular parts of the geometric shapes, spacing responses apart in time by one second or more and changing from one key to the other after a given number of presses. This did not seem to produce more points than if I pressed them randomly. So, after experimenting with the aforesaid ideas I continued to press randomly.
	Subject PB, Experiment 3: I pressed the left and right keys equally often beginning at two (presses) on the right and two on the left, up to 24 on each key.
Undermatching	Subject JP, Experiment 2: For the "+" (VI 10 s) and the "o" (VI 20 s) schedules you could press these mostly continually and be fairly sure of a good score. Of the remaining shapes it was difficult to gain points so I would change to the left-hand key (VI 50 s) for about 40% of the time while these shapes were lit up.
Matching	Subject NL, Experiment 3: For each combination I pressed the button which I felt would increment me most, at the same time occasionally pressing the other button. The number of times that I pressed the lower increment button depended on my assessment of its increment value (see Table 2). The button giving me the most money received in return the greater share of my presses.
Overmatching	Subject RE, Experiment 6: I pressed the cross (VI 10 s) and the circle (VI 20 s) for most of the time, very rarely pressing the triangle (VI 50 s). For the others I pressed the triangle for the majority of the time.
Exclusive preference	Subject CM, Experiment 4: As a result of testing different numbers of presses I decided that I could get more points by only pressing the key giving more points. Changing between keys meant lost opportunities to win points on the richer key.

some contained elements of different rule types. In each case, the rater attempted to assess what performance category each subject's rule or combination of rules would give rise to. In order to check the reliability of this method of categorization, an independent sorting of subjects' performance rules was conducted by an independent rater who was acquainted with matching theory but not with these experiments. The independent rater was given the 30 verbal reports coded by number and was asked to assign each report to one of the five performance categories on the basis of the performance that in his view would have occurred were the rules to be followed. The extent of agreement between the two independent sets of categorizations was then analyzed using the kappa statistic (see Siegel & Castellan, 1988), which provides a measure of interrater agreement corrected for agreements expected on the basis of chance. The outcome of these analyses is reported in the General Results section. Subjects' rule categories, performance categories, and generalized matching equation slope values are given in Table 2.

EXPERIMENT 1

In this experiment, the procedure was similar to that of the two-manipulandum experiments conducted by Bradshaw and collaborators, but an attempt was made to eliminate ordinal cues related to the VI schedules comprising Component A by (a) employing individual geometric shapes to signal the VI schedules and (b) presenting the VI schedules, in the first experimental session, in a nonordinal sequence (i.e., not in an ascending order of interval value). An additional schedule (i.e., concurrent VI 50 s VI 500 s) was also included in the present procedure to provide a more reliable test of Equations 1 through 4.

METHOD

Subjects

Five subjects, 3 female (CT, TM, and JT) and 2 male (KW and JB), took part.

Procedure

The operation of each VI schedule, whether in Component A or Component B, was signaled by a corresponding geometric shape projected onto Key A and Key B, respectively. No COD was employed.

RESULTS

In this and subsequent experiments, data from the last three experimental sessions were

Table 2

A comparison of subjects' schedule rankings,[a] performance rule categories,[b] obtained generalized matching equation slope values, and performance categories.[b]

Subject	Schedule ranking	Rule category	Slope	Performance category	Subjects	Schedule ranking	Rule category	Slope	Performance category
	Experiment 1					Experiment 4			
CT	1,2,3,4=6,5	U	0.53	U	PH	1,2,6,3,4=5	M	1.13	M
TM	1,3,2=6,4,5	I	0.03	I	RC	1,2,3,4,5,6	M	0.92	M
JT	1,2,3,4=6,5	I	0.04	I	DJ	1,2,3,4=5=6	U	0.38	U
KW	1,2,3,5,4=6	I	0.07	I	DK	1,2=3=4=5,6	I	0.03	I
JB	1,2,3,6,5,4	I	−0.06	I	CM	1,2,3,4,5,6	EP	—[c]	EP
	Experiment 2					Experiment 5			
JP	1,2,3,4=5=6	U	0.62	U	DE	1,2,3,4,5,6	O	2.01	O
CP	1,2,3,4=5=6	U	0.59	U	SM	1,2,3,4,5,6	O	1.56	O
DF	1,2,3,6,4,5	U	0.45	U	TJ	1,2,3,4,5,6	M	0.87	M
KW	1,2,3,5,6,4	U	0.37	U	MW	1,2,3=4,5,6	U	0.54	U
IJ	1,2,3,5,6,4	I	0.15	I	JC	1,2,3,4,5,6	I	0.17	I
	Experiment 3					Experiment 6			
NL	1,2,3,4,5,6	M	0.98	M	RE	1,2,3,4,5=6	O	1.81	O
SP	1,2,3,4,5,6	M	0.79	M	BL	1,2,3=4,5=6	M	0.84	M
SC	1,2,3,=4=5,6	U	0.56	U	WJ	1,2,3,4=5=6	U	0.55	U
CL[d]	1,2,3,4,5,6	O	1.12	M	FP	1,2,3,4,5,6	EP	—[c]	EP
PB	1,4,2,5,3,6	I	0.07	I	NW	1,2,3,4,5,6	EP	—[c]	EP

[a] 1 = VI 10 s, 2 = VI 20 s, 3 = VI 50 s, 4 = VI 175 s, 5 = VI 500 s, 6 = VI 720 s; equal preference between pairs of schedules is indicated by =.

[b] I = indifference, U = undermatching, M = approximate matching; O = overmatching, EP = exclusive responding to the richer alternative on all six concurrent schedules.

[c] Linear regression analysis is not possible in these cases because subjects responded on only one alternative of all six concurrent schedules.

[d] Data from Session 16 only.

used to calculate mean response rates and mean frequency of obtained reinforcement on each schedule value. Visual inspection of response rates showed the behavior of all 5 subjects to be stable over the last five sessions.

Responding on Key A and Key B

In Figure 1, responses per minute on Key A and Key B are plotted against reinforcers per hour obtained in Component A. Curves having the forms defined by Equations 1 and 2 were fitted to the data by nonlinear regression analysis (Bradshaw, 1977; Wilkinson, 1961). Derived values of K, r_0, and p^2 (where p^2 is the proportion of the y values that can be accounted for in terms of x, in a curvilinear function; see Bradshaw, Szabadi, & Bevan, 1979; Lewis, 1960) are presented in Table 3. The best fit of Equation 1 to the data was obtained for Subject CT, although p^2 was below the criterion of 80% and r_0 was negative. Because r_0 has been defined theoretically as "the reinforcement frequency needed to maintain half-maximal response rate" (Bradshaw et al., 1976) or, alternatively, "all extraneous alternative sources of reinforcement in the experimental setting" (Herrnstein, 1970), negative values of r_0 become psychologically meaningless. In the case of the remaining 4 subjects, p^2 values were so low as to preclude meaningful interpretation of the derived parameters.

An adequate fit to Equation 2 ($p^2 > 80\%$) was obtained only for Subjects CT and TM (see Table 3). In summary, with respect to responding on both Key A and Key B, Equations 1 and 2 described the data of at best only 1 of the 5 subjects (CT) in this experiment.

The Matching Law

Even when p^2 values are disregarded, Table 3 shows that none of the subjects had similar values of K and r_0 across Equations 1 and 2. Because these parameters are not invariant,

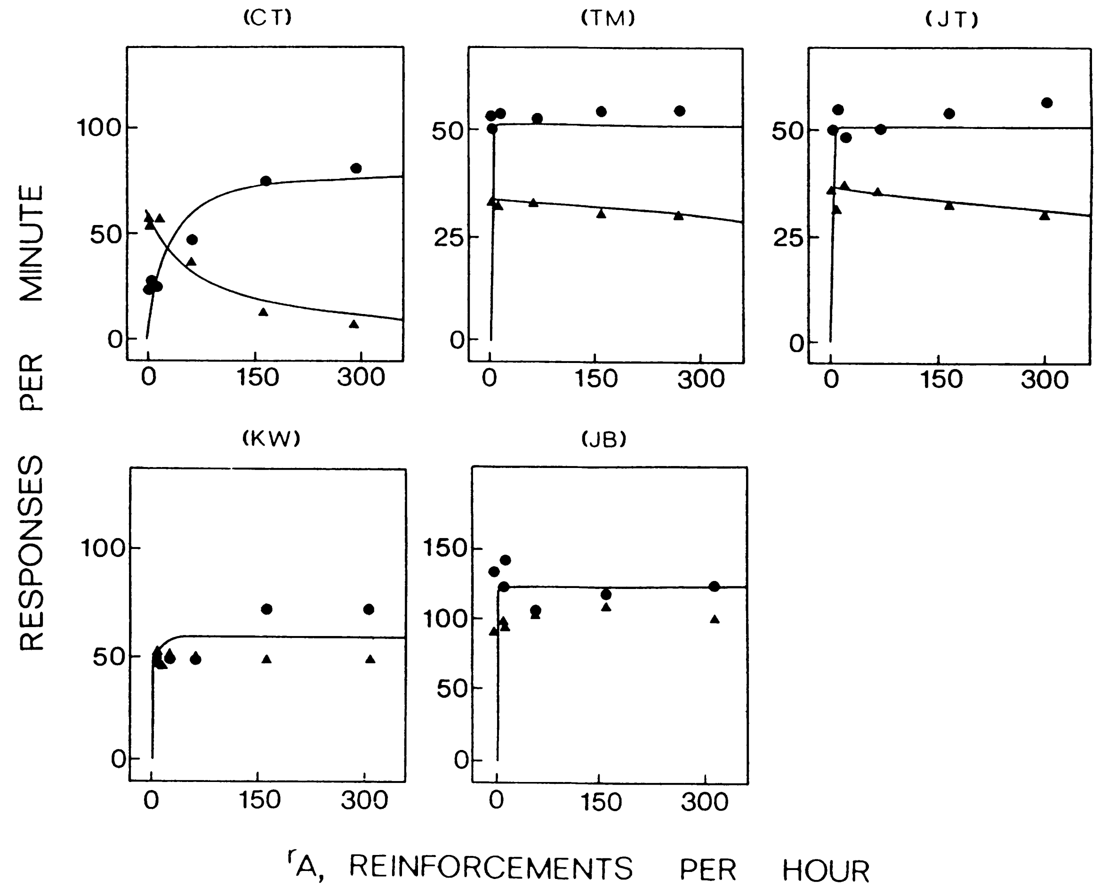

Fig. 1. Mean response rates in Component A (R_A, filled circles) and Component B (R_B, filled triangles) plotted against mean delivered reinforcement frequency in Component A (r_A), for each subject in Experiment 1. Unless otherwise indicated, in this and subsequent figures, data points are based on the last three sessions. Curves are best fit rectangular hyperbolae, fitted by nonlinear regression analysis.

and indeed the differences are great across the two Components A and B, Equations 1 and 2 may not be combined to yield Herrnstein's matching law (i.e., Equation 3).

Figure 2 shows, for each subject, the ratios of response rates in the two components (R_A/R_B) as a function of the ratios of reinforcement frequencies delivered in the two components (r_A/r_B), using double logarithmic coordinates (cf. Baum, 1974). Linear functions were derived by the method of least squares. The slope of the fitted line is a measure of the exponent a, and the intercept is a measure of k, in Equation 4. The behavior of none of the 5 subjects conformed to Herrnstein's matching law (Equation 3) and only 1, CT, might be said to conform to the generalized matching law (Equation 4); even this subject undermatched considerably ($a = 0.53$). The remaining 4 subjects had values of a close to zero.

Verbal Reports

In Experiment 1, Subjects TM, JT, KW, and JB reported schedule preferences that did not accurately reflect the scheduled contingencies (see Table 2). Moreover, as is illustrated in Table 1, they articulated indifference performance rules and, consistent with these rules, their choice performances yielded generalized matching equation slope values close to zero.

Subject CT accurately ranked the three richest schedules with respect to reinforcement frequency, and her performance rules were classified as undermatching. The categorization of her performance rules was consistent

Table 3

Estimated values of the parameters K (response rate at asymptote), and r_0 (a measure of unscheduled reinforcement in the experimental setting) obtained by nonlinear regression analysis from plots of mean response rate versus mean delivered reinforcement frequency, together with p^2 estimates (percentage of variance accounted for by the fitted hyperbola), for all subjects in Experiments 1 through 6. Mean response rates and mean delivered reinforcement frequencies were derived from data obtained in the last three experimental sessions.

	Parameter estimates					
	Upward curve			Downward curve		
Subjects	K	r_0	p^2	K	r_0	p^2
Experiment 1						
CT	82.06	−34.81	76.03	76.41	15.34	95.83
TM	54.32	−62.57	54.00	1,497.10	2,798.86	78.80
JT	53.58	−68.33	11.32	1,028.60	1,890.46	45.75
KW	58.70	−59.97	34.95	—[a]	—	—
JB	124.36	−62.31	13.27	—	—	—
Experiment 2						
JP	27.48	−33.97	20.99	7.59	−12.72	94.00
CP	315.40	−24.26	84.08	—	—	—
DF	229.13	−46.67	20.99	313.87	55.40	75.05
KW	170.94	−47.67	32.10	295.84	76.88	87.03
IJ	161.13	−39.84	0.45	998.44	442.36	43.55
Experiment 3						
NL	343.02	−11.79	95.01	315.05	5.23	95.47
SP	217.69	−16.54	98.11	90.85	−19.89	93.38
SC	141.45	−45.03	92.45	225.87	26.88	96.16
CL[b]	234.04	−58.39	2.96	—	—	—
CL	310.50	−26.60	80.78	252.34	4.25	93.42
PB	146.66	−59.84	39.76	—	—	—
Experiment 4						
PH	288.84	28.81	98.59	178.15	−10.08	95.49
RC	332.64	22.49	94.89	233.42	1.86	95.52
DJ	189.34	−45.77	68.34	519.85	113.36	84.42
DK	87.64	−50.24	16.15	—	—	—
CM	224.28	−14.87	98.08	—	—	—
Experiment 5						
DE	370.24	26.44	96.74	—	—	—
SM	329.24	−5.26	94.31	—	—	—
TJ	278.45	−19.01	81.39	—	—	—
MW	97.30	−51.45	10.63	—	—	—
JC	22.57	−53.74	71.47	457.72	988.04	79.61
Experiment 6						
RE	338.63	19.93	98.99	180.75	−14.85	96.44
BL	300.34	2.71	91.53	239.95	1.95	92.12
WJ	192.63	136.89	85.13	231.80	113.27	84.67
FP	303.00	31.63	98.97	113.16	−21.92	98.61
NW	279.25	−27.00	99.99	—	—	—

[a] Dash indicates that a fit to the data was not possible.

[b] In the case of Subject CL, Experiment 3, two sets of data are presented; the first is derived from the mean of Sessions 13, 14 and 15, and the second is from Session 16 only.

with her generalized matching equation slope value (a = 0.53).

Discussion

When ordinal cues were minimized, the performance of 4 out of 5 subjects on a multiple concurrent VI VI procedure failed to conform to both Equations 1 and 2 or to the generalized matching law. Moreover, the performance of none of the 5 subjects could be described by both Equations 1 and 2, or by Equation 3. These findings differ greatly from those of the

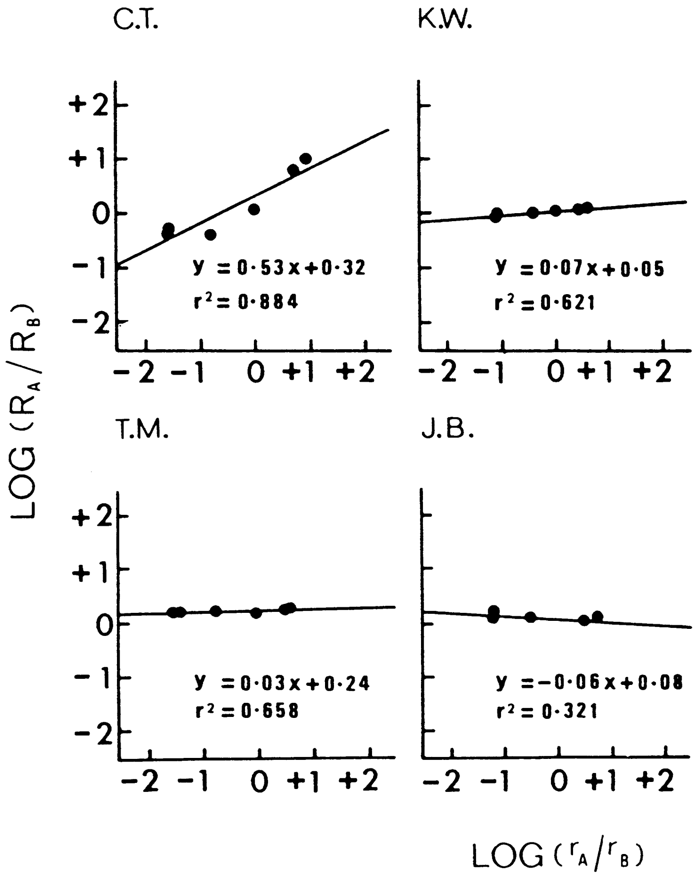

Fig. 2. Ratios of mean response rates in the two components (R_A/R_B) plotted against ratios of mean delivered reinforcement frequencies in the two components (r_A/r_B), using double logarithmic coordinates, for each subject in Experiment 1. Means are based on the last three sessions. The straight lines are the best fit linear functions obtained by the least squares method. Equations for the linear functions are shown in each panel, together with r^2 values (proportions of data variance accounted for by the linear functions).

Bradshaw experiments but are consistent with other studies of human performance on concurrent schedules (Navarick & Chellsen, 1983; Oscar-Berman et al., 1980; Pierce et al., 1981; Poppen, 1982; Schmitt, 1974; Takahashi & Iwamoto, 1986). All 5 subjects' choice performances were consistent with their performance rules.

EXPERIMENT 2

Experiment 1 showed that when ordinal cues were eliminated from the Bradshaw procedure, 4 out of 5 subjects failed to show any evidence of matching. It is possible, however, that the basic procedure could have yielded matching if other variables were altered. For example, although no COD was used in the Bradshaw studies, it has been argued that a COD should generally be included in concurrent schedule procedures (Catania, 1966; McSweeney et al., 1983). This experiment investigated the addition of a COD to the procedure employed in Experiment 1.

METHOD

Subjects

Five subjects, 3 female (DF, KW, and JP) and 2 male (CP and IJ), took part.

Procedure

The apparatus and procedure were identical to those of Experiment 1, except that a 3-s COD was employed throughout.

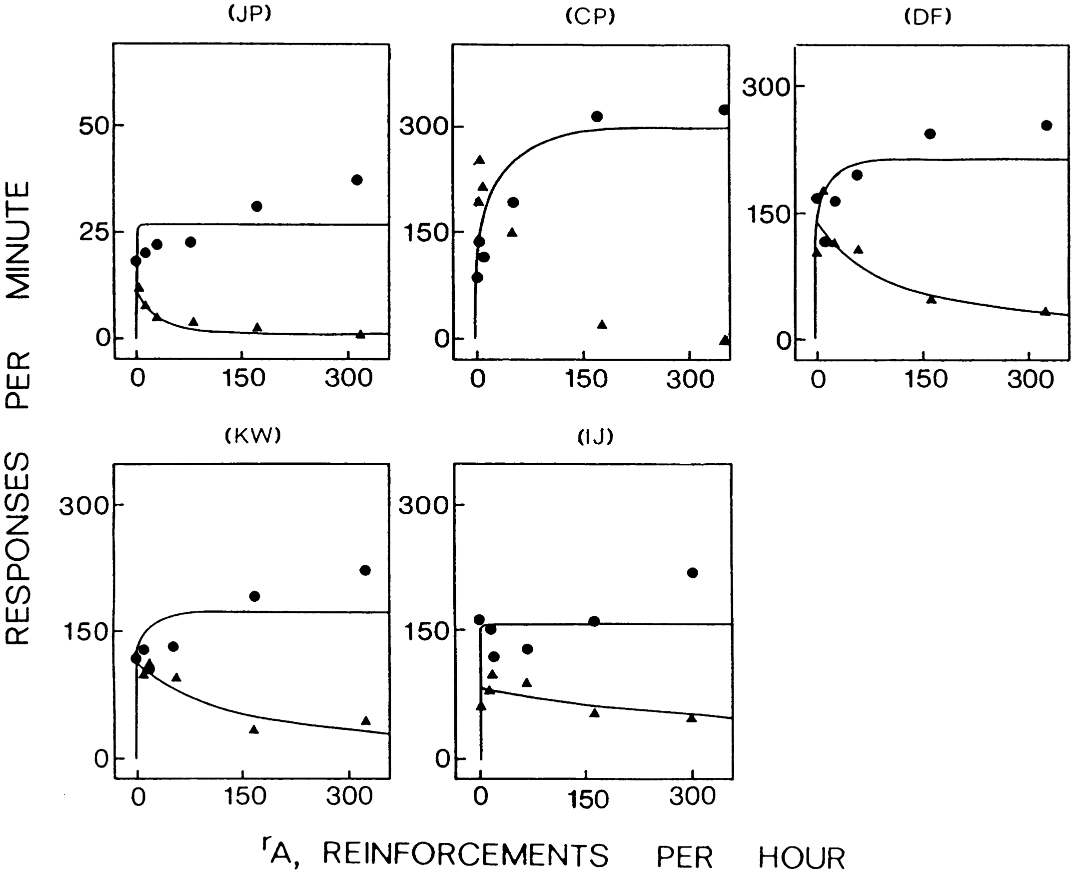

Fig. 3. Mean response rates in Component A (R_A, filled circles) and Component B (R_B, filled triangles) plotted against mean delivered reinforcement frequency in Component A (r_A), for each subject in Experiment 2.

Results

Visual inspection of response rates showed that the behavior of Subjects KW and IJ over the last five sessions was stable. For Subjects JP, CP, and DF, there was a decreasing trend in sensitivity to relative reinforcement.

Responding on Key A and Key B

Figure 3 shows responses per minute on Key A and Key B plotted against reinforcements per hour obtained in Component A. The performance of only 1 of the 5 subjects (CP) could be described by Equation 1, although the r_0 value for this subject was negative (see Table 3). For the remaining subjects, all p^2 values were very low and all derived values of r_0 were negative. An adequate fit of Equation 2 to the data was obtained for only 2 of the 5 subjects (JP and KW).

The Matching Law

There was no case in the performance of any of the 5 subjects in which both K and r_0 were invariant across Equations 1 and 2 (see Table 3). Consequently, these equations may not be combined to yield Equation 3. The regression analyses (Figure 4) show that the data for Subjects JP and CP were well described by Equation 4, although the values of a indicated considerable undermatching. Moreover, Subject CP showed exclusive responding (i.e., he responded only on the richer of the two components) on concurrent VI 50 s VI 10 s, and almost exclusive responding on concurrent VI 50 s VI 20 s; consequently, only five data points could be computed for this subject. Similarly, Subject JP's performance was close to exclusive preference on concurrent VI 50 s VI 10 s. The r^2 values for Subjects DF and

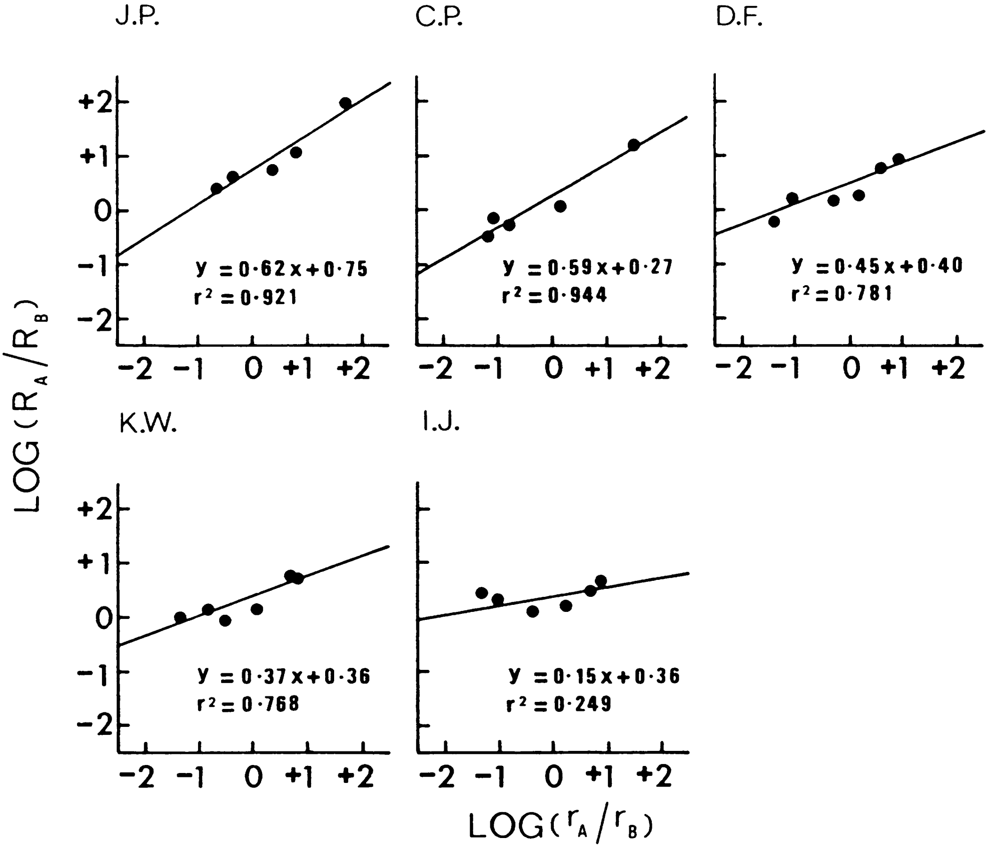

Fig. 4. Ratios of mean response rates in the two components (R_A/R_B) plotted against ratios of mean reinforcement frequencies in the two components (r_A/r_B), using double logarithmic coordinates, for each subject in Experiment 2.

KW were low, and the values of *a* represent gross undermatching. The U-shaped distribution of the data points for IJ (r^2 = 25%; *a* = 0.15) is clearly inappropriate for linear regression analysis.

Verbal Reports

Subjects JP (see Table 1), CP, DF, and KW all reported undermatching performance rules (see Table 2) that were consistent with their undermatching performances. Subjects JP and CP ranked the three richer schedules accurately but reported equal preference among the three leaner schedules. Although JP expressed no preference between the three leaner schedules, there was a slight progressive shift in her relative responding from Component A to Component B as scheduled reinforcement frequency decreased on Component A from VI 175 s to VI 720 s (see Figure 4). For Subjects DF and KW, even though their ranking did not correspond closely to the scheduled contingencies, there was a close correspondence between their stated order of preference and their relative responding on the keys (see Table 2 and Figure 4). Subject IJ ranked only the three richer schedules accurately and reported indifference performance rules consistent with a generalized matching equation slope of 0.15.

Discussion

The addition of a COD to our procedure did little to produce conformity to the generalized matching law of a kind observed in nonhuman studies; the predominant outcome was gross undermatching, as compared to indifference in Experiment 1. The performance of none of these subjects conformed (p^2 > 80%) to both single-rate equations.

An interesting feature was the strong tendency towards exclusive responding shown by 2 subjects on the richest schedules. Given the failure to sample one of the concurrently available alternatives, such responding represents at best trivial matching. Because exclusive responding did not occur in Experiment 1, it may have been the case that the 3-s COD was responsible for its occurrence in Experiment 2.

As was the case in Experiment 1, subjects reported idiosyncratic formulations of the contingencies and response strategies. However, there were fewer instances of alternation strategies; the operation of the COD would have made such strategies particularly unproductive. As in Experiment 1, there was a correspondence between what many of the subjects said and what they did, although this was not invariably true. For example, although Subject JP reported no differential rule for responding on VI 175 s, VI 500 s, and VI 720 s, her relative rates of responding showed some sensitivity to reinforcement frequency on these schedules.

EXPERIMENT 3

One reason the findings of Experiments 1 and 2 were not in agreement with those of Bradshaw and colleagues might have been that our experiments lacked the ordinal cues that operated implicitly in the Bradshaw procedure. Bradshaw and colleagues employed an array of five lights, illumination of one of which corresponded to the operation of a particular VI schedule. In the present experiment, therefore, a stimulus analog of this array was employed in place of the geometric stimuli used in Experiments 1 and 2.

Method

Subjects

Five subjects, 3 female (SP, CL, and PB) and 2 male (NL and SC), took part.

Apparatus and Procedure

The procedure was the same as that used in the first two experiments, except that the schedule stimuli (geometric shapes) used in Experiments 1 and 2 were replaced by a row of six filled circles, one of which was colored red and the remainder were black. The position of the red circle was ordinally related to the scheduled reinforcement frequency in the same manner as the position of the illuminated light in the array employed by Bradshaw and colleagues. When the VI 10-s schedule was operative, the left-most circle was red; when the VI 20-s schedule was operative, the next to left-most circle was red, and so on. The instructions described in the General Method section were modified by replacing references to "shapes" with "row of dots." As in Experiment 1, no COD was employed.

Results

Visual inspection of response rates over the last five sessions showed that the behavior of Subjects NL and PB was stable. This was also true for Subject CL, with the exception of data obtained in her last session. For Subjects SP and SC, there was a slight decreasing trend in sensitivity to relative reinforcement.

Responding on Key A and Key B

Figure 5 shows responses per minute on Key A and Key B plotted against reinforcements per hour obtained in Component A. With the exception of Subject CL, these results are based upon the last three sessions of responding; the function for CL excludes data obtained on the last day of the study and represents the mean of the preceding three sessions. (The reason for this omission is discussed below.) These data, together with those in Table 3, show that the performance of Subjects NL, SP and SC was well described by Equation 1, although in all cases the parameter r_0 was negative; the data for PB and CL were poorly described by Equation 1. Equation 2 provided a good fit to the data for Subjects NL, SP, and SC, with the qualification that r_0 was negative for SP's data. In the case of PB and CL, it was not possible to fit Equation 2 to the data. Disregarding the incidence of negative r_0 values, the single-rate equations provided a very good mathematical fit to the data of 3 subjects; the p^2 values exceeded 92% with a mean of 96%, an outcome similar to that reported by Bradshaw and colleagues.

The Matching Law

There was no case in the performance of any of the 5 subjects in which both K and r_0 were invariant across Equations 1 and 2 (see Table 3). Consequently, these equations may not be combined to yield Equation 3. The regression analyses for the 5 subjects are shown in Figure 6. Except in the case of Subject CL, these results are based upon the last three ses-

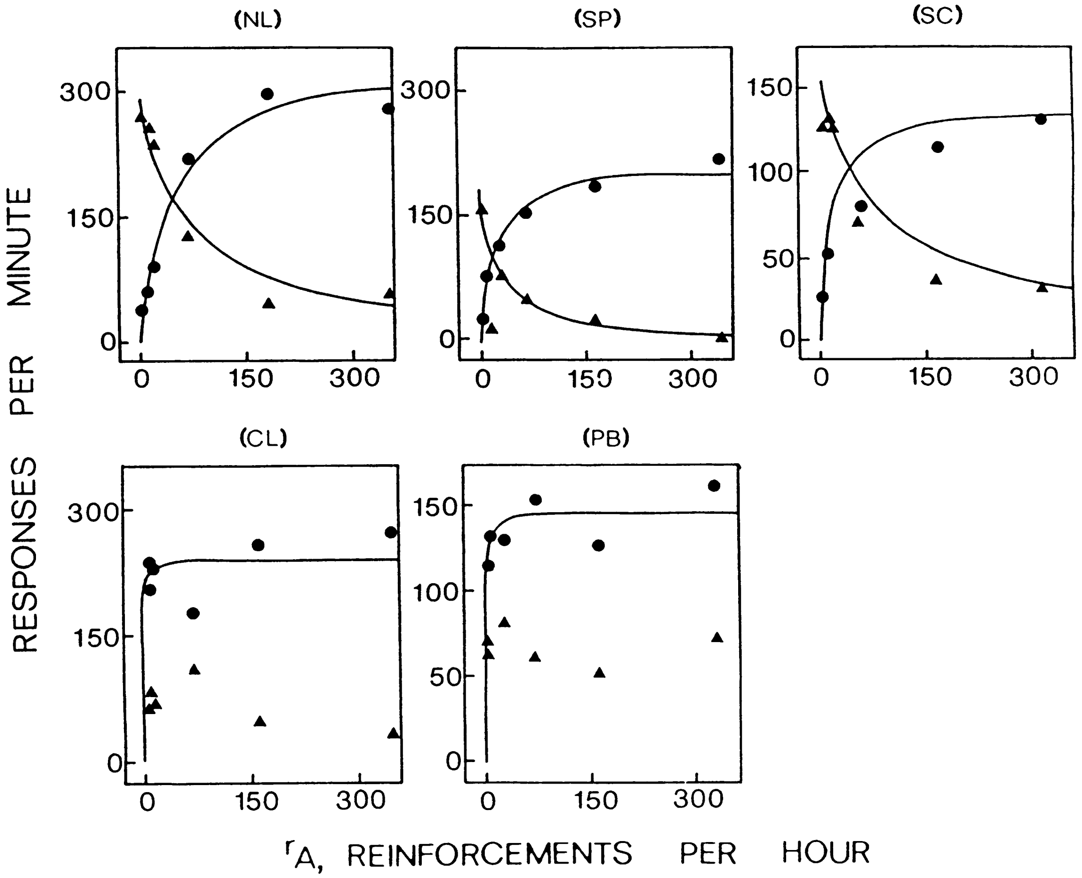

Fig. 5. Mean response rates in Component A (R_A, filled circles) and Component B (R_B, filled triangles) plotted against mean delivered reinforcement frequency in Component A (r_A), for each subject in Experiment 3. With the exception of CL (see text), for whom data points are based on Sessions 13, 14, and 15, means are from the last three sessions.

sions of responding; the function for CL excludes data obtained on the last day of the study and represents the mean of the preceding three sessions. The data for 3 of the 5 subjects (NL, SP, and SC) were well described by the generalized matching equation, and matching-like responding was present from the second session onwards; by the third session, values of *a* were 1.7, 1.3, and 0.8, respectively. This was not true for the remaining 2 subjects (CL and PB), each of whose performance was poorly described by the generalized matching equation, with values of *a* indicative of indifference.

Subject CL

In her last experimental session, the performance of Subject CL differed markedly from that in the preceding sessions. Figure 7 shows responses per minute on Key A and Key B plotted against reinforcements per hour obtained in Component A during the last session. These data were adequately described by Equation 1 and better described by Equation 2. In terms of the generalized matching equation, the regression analysis shows that these data were well described, with a value of *a* close to ideal matching.

Verbal Reports

Subject NL accurately ranked the schedules and reported performance rules that were categorized as matching (see Table 1); his performance showed ideal matching. Although for much of the study Subject CL's performance had been insensitive to reinforcement frequency on the leaner concurrent VI schedules, it changed markedly from Session 15 to 16. She revealed at the end of the experiment that she had spent some time during the evening before her last session with Subject NL; the

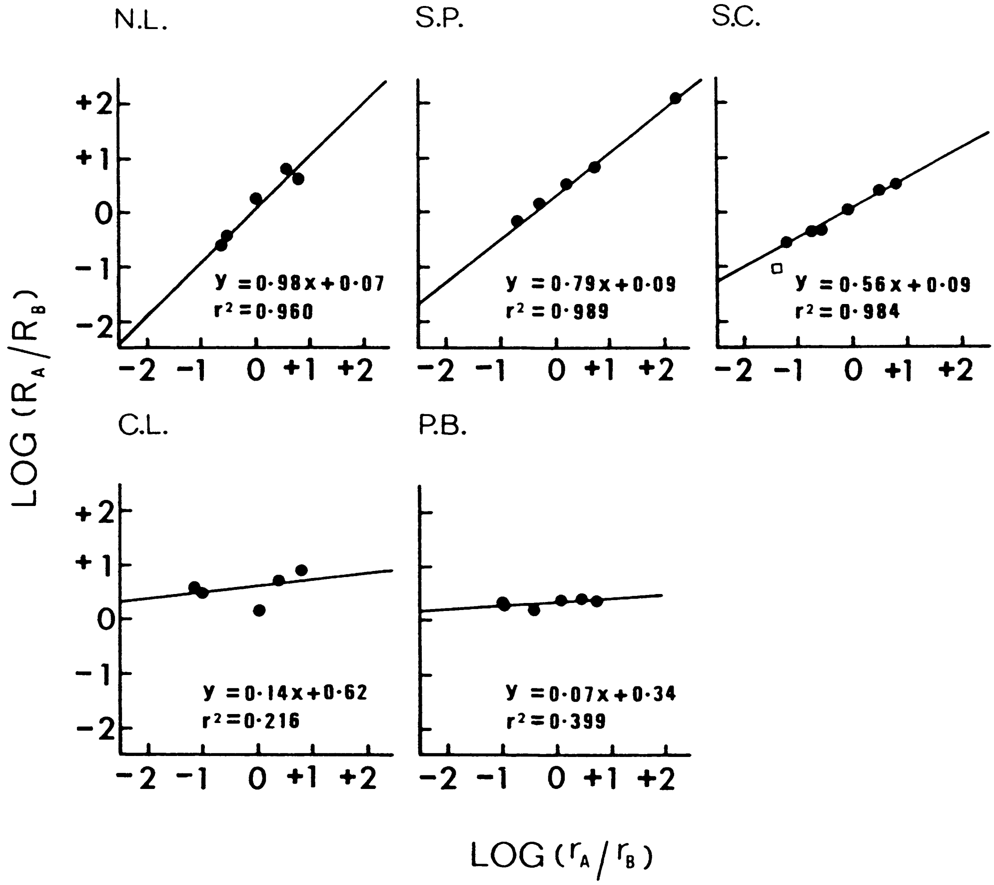

Fig. 6. Ratios of mean response rates in the two components (R_A/R_B) plotted against ratios of mean delivered reinforcement frequencies in the two components (r_A/r_B), using double logarithmic coordinates, for each subject in Experiment 3. With the exception of CL, for whom data points were based on Sessions 13, 14, and 15, means are from the last three sessions.

following day, her performance in Session 16 was close to ideal matching. She reported: "In the last session I cheated! Last night I was told by [NL] to press the left-hand key quickly and frequently when the red dot [on the right-hand key] was in the low scoring position in order to gain more points." This new rule for responding on the leaner schedules contrasted markedly with that which she said she had formulated prior to her conversation with NL, which was a combination of an overmatching rule on some schedules and an antimatching rule on the leaner schedules. For the last session she accurately ranked the schedules, and her performance rules were categorized as overmatching, although in fact her performance was closer to ideal matching (see Table 2).

Subject SP ranked the schedules accurately and reported an approximate matching rule consistent with her performance. Subject SC accurately ranked the three richer schedules but expressed an equal preference for VI 50 s, VI 175 s, and VI 500 s. Her undermatching performance rule was consistent with her schedule behavior (see Table 2). Subject PB's ranking of the stimuli diverged greatly from the scheduled reinforcement frequencies. She reported an indifference rule (see Table 1) consistent with her generalized matching equation slope of 0.07.

Discussion

By the end of this experiment, 4 subjects showed matching-like behavior, with more than 96% of data variance (r^2) accounted for by the generalized matching equation and performances similar to those observed in studies

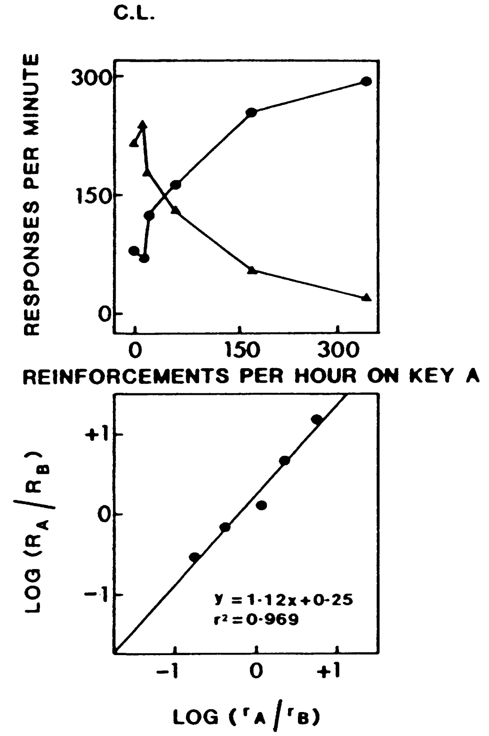

Fig. 7. Absolute response rates in Component A (R_A, filled circles) and Component B (R_B, filled triangles) plotted against delivered reinforcement frequency in Component A (r_A), for Subject CL in Session 16, are presented in the upper panel. The best fit linear function of response ratios over reinforcer ratios is presented in the lower panel.

with nonhumans. This was in contrast to Experiment 1, in which ordinal cues were not employed, and in which the only subject to approximate a matching-like performance (CT) showed considerable undermatching. Two aspects of the development of schedule performances in this experiment were particularly interesting. First, with the exception of CL, matching-like behavior, when it did occur, developed rapidly over the first few sessions and prevailed until the end of the experiment. Second, in the case of CL, an indifference performance in earlier sessions changed abruptly to one of ideal matching in the final session. This sudden change was consistent with her report that following her conversation with Subject NL, she adopted his response strategy. These two features, namely rapid convergence on a matching-like strategy and abrupt qualitative change in the pattern of behavior allocation caused by another subject's verbal instruction, suggest that much of the schedule performance of these subjects may be rule governed.

It appears that the ordinal cues employed in this experiment exerted their effect by virtue of the fact that they functioned as verbal stimuli (see General Discussion), providing subjects with implicit information about the scheduled relative reinforcement frequencies comprising Components A and B. Subjects who interpreted the cues in this fashion had only to sample the schedules to corroborate the implicit information. The ordinal schedule cues appeared to reduce the probability of subjects' formulating complex numerical response alternation rules, such as were reported by those subjects in Experiments 1 and 2 whose behavior was insensitive to the scheduled contingencies.

EXPERIMENT 4

In Experiment 3, which employed "implicit" ordinal schedule stimuli, 4 out of 5 subjects showed conformity to the generalized matching equation. The purpose of the present experiment was to determine whether the addition of a 3-s COD to this procedure might affect the incidence of matching.

METHOD

Subjects

Five subjects, 4 female (RC, DJ, DK, and CM) and 1 male (PH), took part.

Apparatus and Procedure

The apparatus and procedure were identical to those employed in Experiment 3, except that a 3-s COD was implemented.

RESULTS

Visual inspection of response rates showed the behavior of Subjects PH, DJ, DK, and CM to be stable over the last five sessions. For Subject RC, there was a slightly decreasing trend in sensitivity to relative reinforcement.

Responding on Key A and Key B

Figure 8 shows that the performance of 3 subjects (PH, RC, and CM) was well described by Equation 1, although r_0 was negative for Subject CM. For the remaining 2 subjects (DJ and DK), Equation 1 provided a poor description of the data. Equation 2 provided a good fit to the data of 3 subjects (PH,

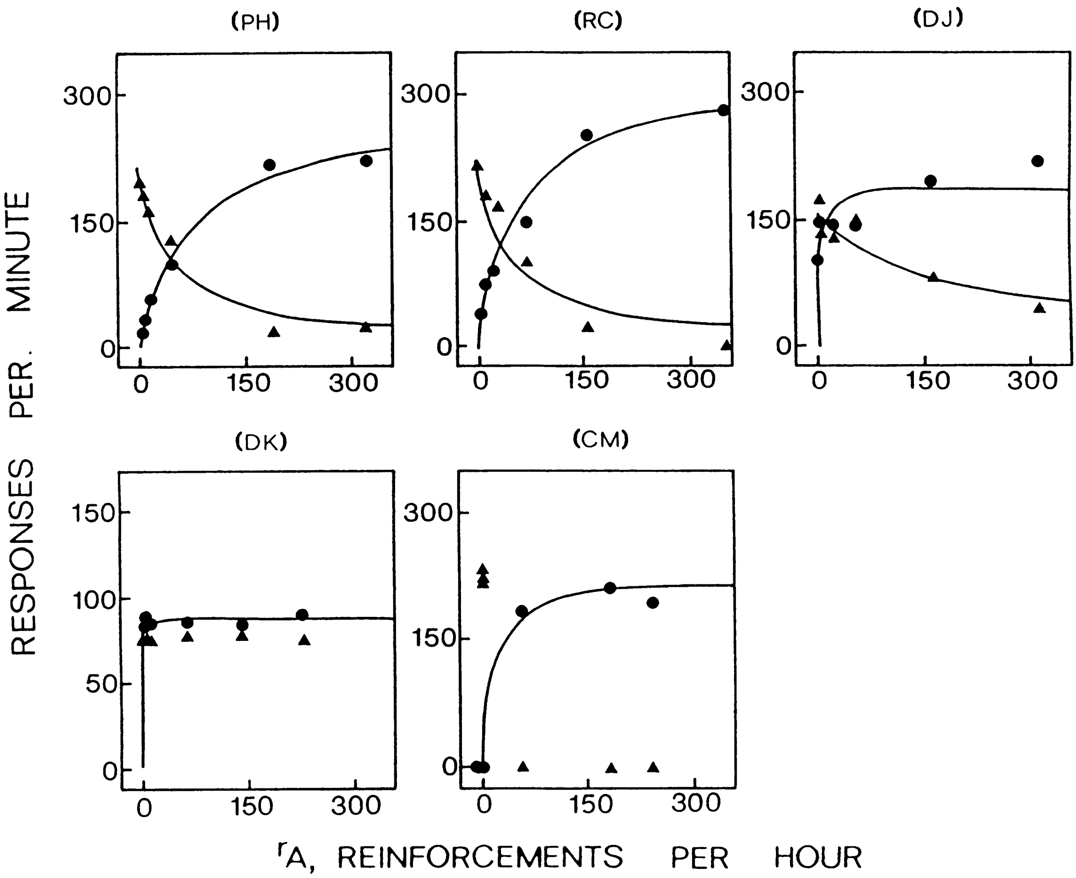

Fig. 8. Mean response rates in Component A (R_A, filled circles) and Component B (R_B, filled triangles) plotted against mean delivered reinforcement frequency in Component A (r_A), for each subject in Experiment 4.

RC, and DJ), although r_0 was negative for PH. For the remaining 2 subjects (DK and CM), it was not possible to fit Equation 2 to the data.

The Matching Law

There was no instance in the performance of the 5 subjects in which both K and r_0 were invariant across Equations 1 and 2 (see Table 3). Consequently, these equations may not be combined to yield Equation 3. The regression analyses presented in Figure 9 show that the data for PH and RC were well described by the generalized matching equation, with values of a close to ideal matching. The performance of DJ represented undermatching, but of a kind rarely encountered in studies of nonhumans, particularly when a COD of 3 s is employed (McSweeney et al., 1983). The generalized matching equation provided a poor description of the data for Subject DK, whose performance could be described as indifference, with a value of a close to zero. In the last two sessions, Subject CM responded exclusively to the richer alternative of all six concurrent schedules; this pattern of behavior is unknown in the concurrent VI literature on nonhumans and cannot be described by the generalized matching equation.

Verbal Reports

PH and RC, both of whom accurately ranked the schedule stimuli, reported approximate matching rules consistent with their schedule performances. Subject DJ accurately ranked the three richer schedules but expressed an equal preference for the three leaner schedules. She reported an undermatching performance rule that was consistent with her schedule performance. DK expressed an equal preference for all but the richest and leanest schedules and reported an indifference rule consisting of

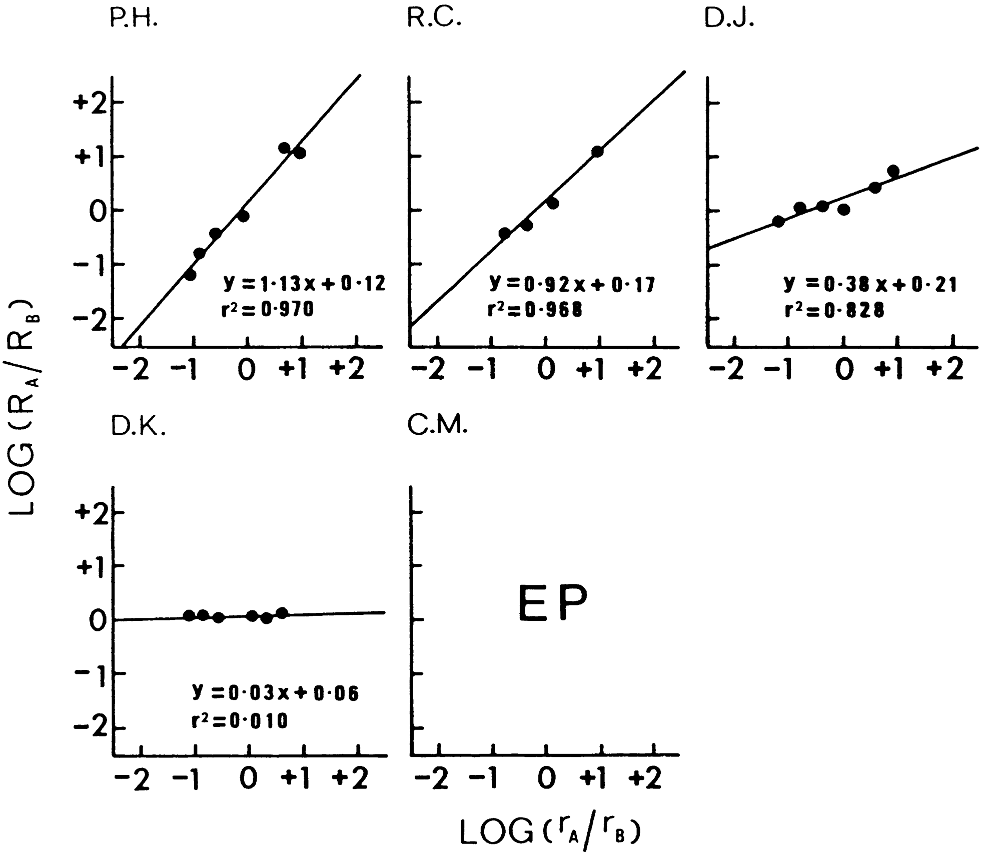

Fig. 9. Ratios of mean response rates in the two components (R_A/R_B) plotted against ratios of mean delivered reinforcement frequencies in the two components (r_A/r_B), using double logarithmic coordinates, for each subject in Experiment 4. Exclusive responding to the richer alternative of all schedules precluded linear regression analysis and is denoted by EP.

mathematically determined alternating patterns of responding across each pair of schedules. This was consistent with an indifference schedule performance with a slope of 0.03.

Subject CM accurately ranked the schedules but reported an exclusive preference performance rule (see Table 1). This rule persisted despite the fact that in earlier sessions (i.e., 3 through 9) before exclusive responding to all six concurrent schedules developed (see Table 2), her session score averaged 9 points higher than in the final sessions when exclusive responding was firmly established.

Discussion

These results, when compared with those of Experiment 3, suggest that a COD contingency does not enhance the effect of ordinal schedule stimuli in such a way as to produce a reliable convergence on ideal or perfect matching. The additional contingency appears to complicate subjects' assessment of the contingencies, resulting in the delayed onset of matching or the emergence of idiosyncratic strategies such as responding exclusively to the richer component of all six concurrent schedules. In the case of exclusive preference, it could be argued that although responding is confined to one alternative, the empirical matching relation is still upheld. This would, however, certainly be the most trivial form of matching (Timberlake, 1982) and, furthermore, is inconsistent with Herrnstein's (1974) original stipulations concerning the relevance of the empirical matching equation.

The 2 subjects whose performance best approximated matching reported performance rules that were consistent with this outcome. Additional evidence for the role of verbal behavior comes from Subject DK. She reported devising complex mathematical rules for alternating responding between the keys, and despite the presence of a COD contingency that heavily penalized frequent alternation, her performance (indifference) was entirely consistent with this account.

EXPERIMENT 5

The schedule stimuli employed in Experiments 3 and 4 were somewhat ambiguous ordinal cues and did not reliably occasion matching behavior. In this experiment, these stimuli were placed with instructions explicitly defining the ordinal relations between schedules with respect to reinforcement frequency. The geometric shapes employed in Experiments 1 and 2 were reinstated.

METHOD

Subjects

Five subjects, 2 female (SM and JC) and 3 male (DE, TJ, and MW), took part.

Apparatus and Procedure

The apparatus and procedure were identical to those employed in Experiment 1 (no COD was used), except that the instruction for the second and subsequent sessions contained the following addition:

> You will find that more points are available when the keys are lit up with certain shapes. This relationship can be described in the following way: + > ○ > △ > — > ⊙ > □ where ">" means "provides more points than" and +, ○, △, —, ⊙, □ are the six different key shapes.

RESULTS

Visual inspection of the response rates showed the behavior of all 5 subjects to be stable over the last five sessions.

Responding on Key A and Key B

Figure 10, together with Table 3, show that Equation 1 provided an adequate fit to the data of 3 of the 5 subjects (DE, SM, and TJ), although r_0 was negative in the latter two cases. Equation 2 could be fitted to the data of only Subject JC.

The Matching Law

The regression analyses presented in Figure 11 show that in terms of r^2 estimates, the performances of 3 subjects (DE, SM, and TJ) were well described by the generalized matching equation, although DE and SM overmatched to an extent not encountered in the literature on nonhumans. Both DE and SM overmatched on the first session (a = 1.30 and a = 1.40, respectively), and this became more extreme as the experiment progressed. TJ produced a good approximation to matching in the first concurrent VI VI session (a = 0.80) and thereafter deviated little from ideal matching. MW also showed matching-like behavior (a = 0.70) on the first concurrent VI VI session and ideal matching in Session 4; thereafter performance drifted to undermatching. The performance of JC could be described as gross undermatching or indifference throughout the experiment.

Verbal Reports

Each subject accurately ranked the six schedules, the only exception being MW, who ranked VI 50 s and VI 175 s as equal. All subjects reported performance rules consistent with their schedule behavior (see Table 2).

DISCUSSION

For 4 of the 5 subjects, performance very early in the experiment approximated matching. To this extent it appears that explicit instructions about the scheduled relative reinforcement frequencies can, at least in the short term, establish a matching-like pattern of responding. As the experiment progressed, however, the performance of Subjects DE and SM diverged in the direction of gross overmatching, whereas the performance of MW drifted towards undermatching. Clearly, ordinal instructions alone are insufficient to maintain ideal matching, and the scheduled contingencies do not appear to be sufficient to prevent a drift away from good approximation to matching established in early sessions (DE, SM, TJ, and MW).

EXPERIMENT 6

This experiment was conducted to determine the effects of the addition of a COD to the procedure used in Experiment 5.

METHOD

Subjects

Five subjects, 3 female (RE, BL, and FP) and 2 male (WJ and NW), took part.

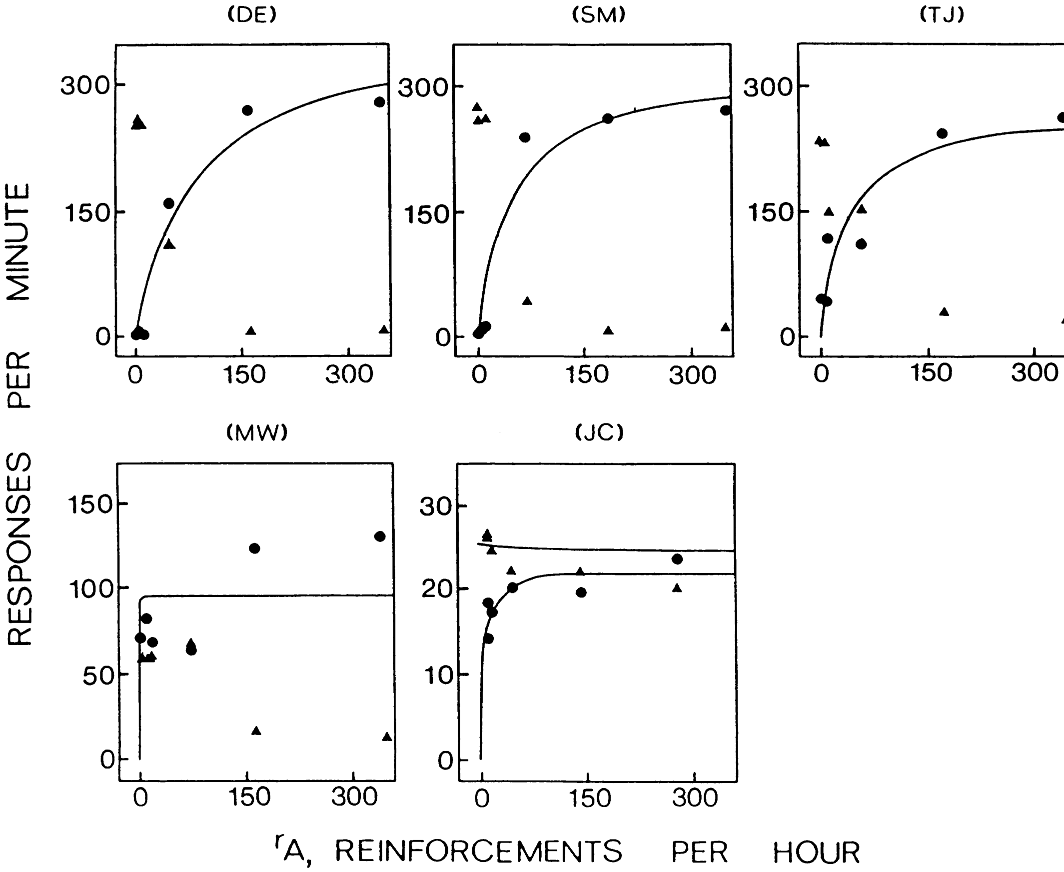

Fig. 10. Mean response rates in Component A (R_A, filled circles) and Component B (R_B, filled triangles) plotted against mean delivered reinforcement frequency in Component A (r_A), for each subject in Experiment 5.

Apparatus and Procedure

The apparatus and procedure were identical to those in Experiment 5, except that a COD of 3 s was employed.

Results

Visual inspection of response rates showed the behavior of Subjects FP and NW to be stable over the last five sessions. For RE and WJ there was a slightly decreasing, and for BL an increasing, trend in sensitivity to relative reinforcement.

Responding on Key A and Key B

Figure 12 shows that the performance of all 5 subjects was well described by Equation 1, with the exception that r_0 was negative for NW (see Table 3). Equation 2 provided a good fit to the data of 4 subjects (RE, BL, WJ, and FP), although r_0 was negative for RE and FP. It was not possible to fit Equation 2 to the data of NW.

The Matching Law

For all 5 subjects, there was no case in which both K and r_0 were invariant across Equations 1 and 2 (see Table 3). Consequently, these equations may not be combined to yield Equation 3. The regression analyses presented in Figure 13 show that in terms of r^2 estimates, the data of 3 subjects (RE, BL and WJ) were well described by the generalized matching equation, although Subject RE showed gross overmatching. Subject BL showed approximate matching, and WJ undermatched. Matching-like behavior was apparent in all subjects in the first session of exposure to the concurrent schedules (RE, a = 0.90; BL, a = 0.50; WJ, a = 0.80; FP, a = 0.90; and NW,

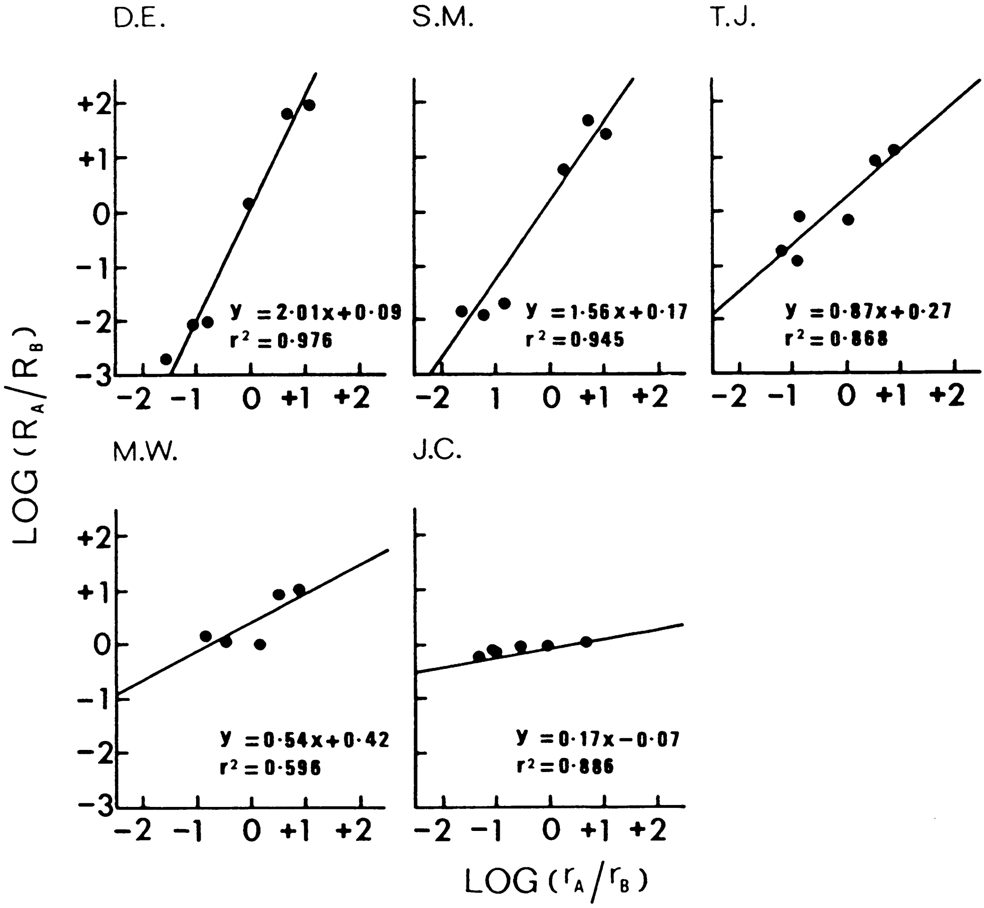

Fig. 11. Ratios of mean response rates in the two components (R_A/R_B) plotted against ratios of mean delivered reinforcement frequencies in the two components (r_A/r_B), using double logarithmic coordinates, for each subject in Experiment 5.

$a = 0.60$). (The slope estimate for NW, however, is based on only two data points because this subject was already responding exclusively on the four remaining concurrent schedules in that session.) By the end of the experiment, both FP and NW responded exclusively to the richer component on all concurrent schedules except VI 50 s VI 50 s, where FP responded equally to both components.

A marked tendency toward exclusive responding was also apparent in the performance of RE, who responded exclusively to the richer component on concurrent VI 50 s VI 10 s and in some sessions on concurrent VI 50 s VI 20 s, and virtually exclusively to the richer component on the three leanest concurrent schedules. From Session 4 to Session 11, Subject BL responded exclusively to the component with the higher reinforcement frequency on the two richest, and sometimes on the two leanest, concurrent schedules.

Verbal Reports

Subjects FP and NW, both of whom correctly ranked the schedules, articulated exclusive preference rules and responded exclusively to the richer alternative on all six schedules. For the remaining subjects, although ranking was broadly accurate, there were some instances of equal preference of schedules with differing reinforcement frequency. Subject RE reported an overmatching rule (see Table 1), and Subjects BL and WJ reported undermatching rules; these were consistent with their schedule performance (see Table 2).

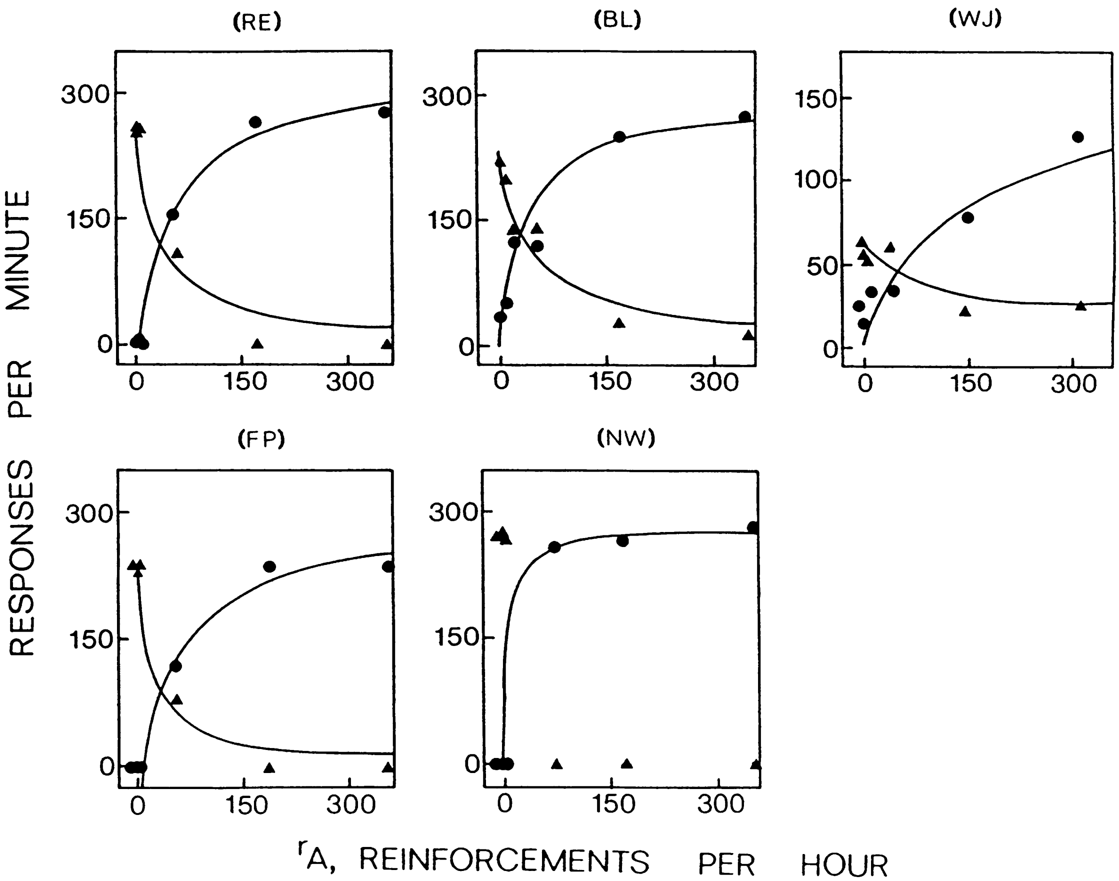

Fig. 12. Mean response rates in Component A (R_A, filled circles) and Component B (R_B, filled triangles) plotted against mean delivered reinforcement frequency in Component A (r_A), for each subject in Experiment 6.

Discussion

This experiment is of interest not only because it documents choice behavior that is peculiar to human subjects but also because it extends the literature on instructed human operant behavior. In complex schedules of this type, a simple verbal instruction relating to the ordinal value of the schedule, as used in Experiments 5 and 6, gives rise to a variety of performances. When compared with the findings of Experiment 5, there was no evidence that the introduction of the COD facilitated the development of matching. Indeed, as was the case in Experiment 4, the COD appears to have given rise to a form of responding on concurrent schedules unique to humans (i.e., exclusive preference) but with very few instances of undermatching or indifference (cf. Experiments 1 and 2).

GENERAL RESULTS

Herrnstein's Single-Rate Equations and the Empirical Matching Relation

If a p^2 estimate equal to or greater than 80% is adopted as the relevant criterion (de Villiers, 1977; de Villiers & Herrnstein, 1976), the data of half the subjects participating in this series of experiments were satisfactorily described by either Equation 1 or Equation 2, but only one third were described by *both* Equations 1 and 2. The single-rate equation does not, therefore, reliably describe human responding in the multiple concurrent VI schedule procedure employed in these experiments. Recent assessments of the literature (Dougan & McSweeney, 1985; Warren-Boulton, Silberberg, Gray, & Ollom, 1985) have shown that much the same is true of nonhuman performance.

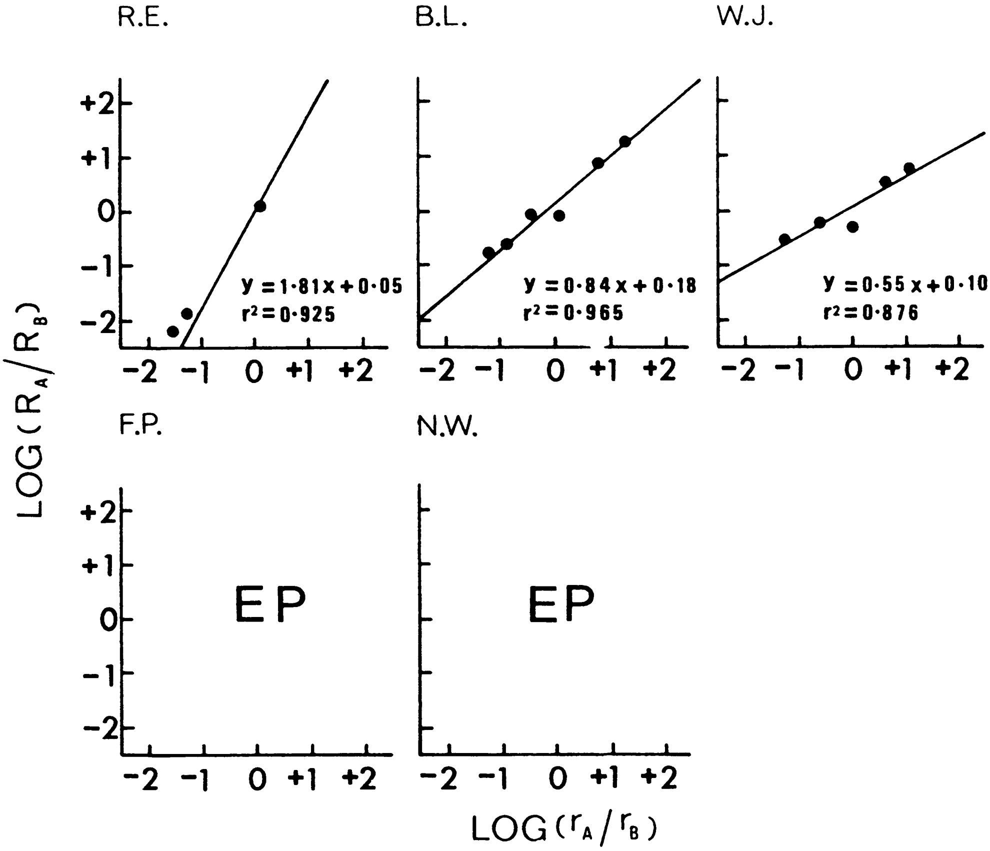

Fig. 13. Ratios of mean response rates in the two components (R_A/R_B) plotted against ratios of mean delivered reinforcement frequencies in the two components (r_A/r_B), using double logarithmic coordinates, for each subject in Experiment 6. Exclusive responding to the richer alternative of all schedules is denoted by EP.

The results of Bradshaw and colleagues, on the other hand, show excellent conformity of human choice data to the single-rate equation (Bradshaw & Szabadi, 1988). However, by the end of Experiment 3, which of the present series of experiments most closely resembled the Bradshaw procedure, the data of 4 of the 5 subjects were well described by the single-rate equations, with a mean p^2 value of 92% for Equation 1 and 96% for Equation 2. In certain circumstances, therefore, and for certain subjects, Herrnstein's single-rate equations can provide a good mathematical description of human choice data. What constitutes such circumstances will be discussed below.

The p^2 criterion provides an estimate only of the mathematical, as opposed to the theoretical, utility of the single-rate equations. Of those cases meeting the p^2 criterion, the derived parameter r_0 was negative in 62.5% of instances for Equation 1 and 33.3% for Equation 2. Furthermore, 7 of the 17 human subjects in the studies by Bradshaw and colleagues produced data with negative r_0 values (SM, 1976; LK, Bradshaw, Szabadi, & Bevan, 1979; B, C, D, E and F, 1981). The high incidence of negative values of this parameter accords well with that observed in studies with nonhumans (McSweeney et al., 1983). Because r_0 theoretically represents the incidence of unscheduled reinforcement in the experimental setting (Herrnstein, 1970) and is central to the assumptions underlying Herrnstein's derivation of the single-rate equation, negative values of this parameter are theoretically troublesome.

In those cases in which the p^2 criterion was met for both Equations 1 and 2, K and r_0 were not comparable across the two equations. The latter may not therefore be combined to derive the empirical matching relation in this series of experiments. Disparity in the K parameter, which denotes response rate at asymptote, could arise in a two-key procedure as a result of topographical consequences of responding at two different locations and could be accommodated as a bias factor such as occurs in the generalized matching equation. The parameter r_0, which denotes unscheduled reinforcement in the setting, may vary in like fashion, but this is less plausible. Unfortunately, Bradshaw and colleagues did not report the values of these parameters for both Equations 1 and 2 in the studies in which they used a changeover-key procedure, which should have minimized discrepancies in K and r_0 attributable to topographical variables. However, the often wide discrepancies observed in the data from the present experiments suggest that both parameters, r_0 in particular, function as free variables accommodating any variance as yet unexplained in the data.

In terms of goodness of fit, the single-rate equations appear to be an unreliable mathematical model of human data, those from the Bradshaw procedure (and see Experiment 3) excepted. As the data from the present experiments in particular show, the equations also entail major theoretical inconsistencies regarding the derived parameters, a finding that has parallels in studies with nonhumans.

The Generalized Matching Equation

Analyses of the literature on choice and the generalized matching equation show that more than 96% of nonhuman subjects' choice performance is characterized by values of a ranging from 0.5 to 1.3 (see Baum, 1979; Horne, 1986; Wearden & Burgess, 1982). The data for the 17 human subjects reported in studies by Bradshaw and colleagues are also characterized by a majority (82%) of a values falling within this range, with two instances falling outside on the overmatching side and one instance of more extreme undermatching. In the six experiments reported here, the performance of only 13 of 30 human subjects (43.3%) fell between these boundaries. Of the remaining subjects, 8 showed indifference (mean value of $a = 0.06$), 3 undermatched (mean value of $a = 0.40$), 3 showed gross overmatching (mean value of $a = 1.79$), and 3 responded exclusively to the richer alternative of all six concurrent schedules (see Table 2).

Verbal Reports: Performance Rules and Matching

Taking all six experiments together, analyses were conducted of the relationship between the rules articulated by subjects in their replies to the questionnaire and their actual performance on the schedules.

Indifference rules. The 8 subjects with rules in this category produced generalized matching equation slope values ranging from −0.06 to 0.17 (see Table 2). Many of these subjects reported that their responding was subject to particular mathematical rules throughout the experiment. Although such reports are retrospective, they furnish good evidence of the efficacy of rule-governed behavior, because it is difficult to claim that subjects first responded and then described post hoc their responses in terms of the specific mathematical sequences they so closely mirrored (see, e.g., PB, Experiment 3).

Undermatching rules. Nine subjects had rules in this category and produced slope values ranging from 0.37 to 0.62. A common feature of undermatching performance rules was that they specified a mixture of approximate matching on some schedules with numerically based alternation strategies on others, and this mixture was reflected in obtained response patterns.

Approximate matching and overmatching rules. The 6 subjects reporting approximate matching rules had slope values ranging from 0.79 to 1.13, whereas the 4 subjects with overmatching rules had slope values ranging from 1.12 to 2.01. A possible reason for the overlap in the approximate matching and overmatching range of slope values is discussed below.

Exclusive preference. Of the 3 subjects reporting exclusive preference rules, 2 (CM and FP) initially responded on both alternatives and thus sampled all the scheduled contingencies. Despite the fact that they scored more points when they sampled both alternatives than when they later adopted the pattern of responding exclusively on the richer alternative, they persisted with the latter behavior. Though the evidence is correlational, the peculiarly maladaptive nature of this uniquely human pattern of responding suggests that it has its origins as a behavioral strategy in rule-governed behavior (see Lowe, 1979; Weiner, 1965). Behavior of this kind has often been encountered in studies of human problem solving (Johnson-Laird & Wason, 1970) and hypothesis testing (Levine, 1971), in which it is found that humans tend to have poor recall of relations between strategies and outcomes or do not integrate these relations rationally.

For all 30 subjects, the categorization (conducted by the first author) of their performance rules was compared with the categorization of their schedule performance (as indicated by slope values; see General Method Section and Table 2); there was a 96.7% agreement between these categorizations. The one instance of a disagreement between performance rule category and outcome was that of Subject CL in Experiment 3. Her performance rules for her last session were assigned to the overmatching category, whereas her schedule performance approximated matching ($a = 1.12$).

In order to validate the above analysis an independent rater, acquainted with matching theory but not with these experiments, conducted an independent categorization of each subject's performance rules, using the same five categories. There were five instances of disagreement between performance rule category and outcome for this rater: Subject RC showed approximate matching but her performance rules were categorized as undermatching; DE and RE overmatched but their performance rules were categorized as approximate matching and, conversely, TJ and BL showed approximate matching although their performance rules were assigned to the overmatching category (see Table 2). The independent validation of the categorization procedure suggests that the attempt to relate rule category to schedule performance has a sound objective basis, except in the case of approximate matching and overmatching categories, for which there was also one discrepant outcome in the first author's attempt at classification. This may have arisen because performance rules within these categories often contained statements that a particular alternative was pressed for "most of the ten minutes" or "lots of the time." Such statements form a poor predictive basis for discriminating between approximate matching and overmatching.

The extent of interrater agreement on the categorization of performance rules was also calculated; there were six instances of disagreement between the two raters, yielding a kappa coefficient of agreement (Siegel & Castellan, 1988) of 0.74 ($p < .01$). These analyses provide good evidence that an independent rater is able to act as a "speaker-listener" (Skinner, 1957) to predict effectively the relation between performance rules and choice behavior.

Verbal Reports:
Schedule Preferences and Matching

We analyzed the relationship between subjects' preferences for the schedules (i.e., the number of schedules accurately ranked) and their matching equation slope values. A Spearman rank order correlation coefficient of 0.70 ($p < .05$) was obtained for these variables and showed that the more accurate the ranking, the higher was the slope of the generalized matching equation. Table 2 shows that, with the exception of JC, none of the 11 subjects who accurately ranked the schedules had slopes lower than 0.79. We also analyzed the relationship between subjects' schedule rankings and response rates on the Component A schedules. The Kendall's coefficient of concordance was 0.63 ($p < .01$), indicating that a positive relationship existed between stated schedule preference and absolute rate of responding.

GENERAL DISCUSSION

The series of experiments reported here is the most extensive yet conducted to investigate both the conditions under which matching may or may not be observed in human behavior and the role of verbal behavior, particularly rules, in human choice. Comparison of our findings with those derived from studies of nonhumans and from those studies of human choice that have employed similar procedures is clearly an important part of this analysis. In our studies, unlike those conducted by Bradshaw and colleagues with humans, less than half the subjects' performances resembled those typically reported in animal choice studies. For many of the remaining subjects, there were not mere "deviations" from the matching typically observed in nonhumans; rather their performance was qualitatively different and could not be described by any of the matching equations.

These findings are consistent with other studies of human behavior on concurrent schedules that have also reported marked departures from matching. Gross undermatching and indifference have been reported for human subjects in studies by Navarick and Chellsen (1983), Oscar-Berman et al. (1980), Pierce et al. (1981), Schmitt (1974), Schroeder (1975), and Takahashi and Iwamoto (1986). Moreover, both Pierce et al. (1981) and Takahashi and Iwamoto (1986) have reported negative values of *a*. The latter authors, for instance, found that in 60% of their subjects' data, *a* ranged from −0.01 to −2.27, with a mean of −0.47, despite the implementation of a 5-s COD. In addition, studies by Schroeder and Holland (1969), Schmitt (1974), and Wurster and Griffiths (1979) reported gross overmatching in the performance of some subjects, although only on some concurrent VI schedule values. Together with the data from our six experiments, these findings clearly demonstrate that human subjects showing ideal matching, or even a close approximation to it, are the exception rather than the rule in the literature (see also Silberberg, Thomas, & Berendzen, 1991). These findings have close parallels in the literature on human operant performance on single schedules, in which differences between humans and nonhumans in both response patterning and sensitivity to schedule parameters have been widely reported (Dugdale & Lowe, 1990; Hayes, Brownstein, Haas, & Greenway, 1986; Hayes, Brownstein, Zettle, Rosenfarb, & Korn, 1986; Hayes, Zettle, & Rosenfarb, 1989; Lowe, 1979, 1983; Matthews et al., 1977; Shimoff, Matthews, & Catania, 1986; Weiner, 1969).

It might be argued, as indeed it often is when differences between nonhuman and human performance are observed in other operant tasks (e.g., Baron, Perone, & Galizio, 1991; Perone, Galizio, & Baron, 1988), that the departure from the typical findings in nonhumans may be due to procedural differences between nonhuman and human studies. This, of course, can never be ruled out, because it is impossible to create exactly the same experimental conditions for different species. For example, it might be argued that too few sessions have been conducted in these and other human studies so that there was insufficient exposure to the scheduled contingencies for behavior to become stable and for matching to emerge. The 16-session multiple concurrent VI schedule procedure used in the six experiments reported here was designed specifically for comparison with the Bradshaw studies. Although relatively few compared with those typical in nonhuman experiments, the number of sessions we employed was the same as in the Bradshaw studies and they far exceed those in most other studies of human choice. In addition, it should be noted that most subjects' performances were stable over the last five sessions and in many cases showed little deviation after the first few sessions, a feature typical of schedule performance in verbally competent human subjects (Lowe, 1979, 1983). It is therefore unlikely that departure from matching in our subjects was due to an insufficient number of sessions. Indeed, in the majority of cases in which there was a trend across the last five sessions, it was away from ideal matching (see, e.g., Experiments 2 and 3).

For similar reasons, it is difficult to maintain that the departures from matching were due simply to the schedules not being discriminable. As in the Bradshaw studies, we separated schedule presentations with 5-min rest intervals in order to minimize schedule interactions, and we used distinctive geometric stimuli to signal the schedules. The most compelling evidence against the discriminability argument, however, comes from (a) the observation that increased exposure to the schedules often led to increased departure from matching and (b) the behavior of those subjects (e.g., JC, CM, FP, and NW) who ranked the schedules with complete accuracy but failed to show matching, responding on the schedules in a manner unprecedented in the literature on nonhumans.

It is sometimes suggested (e.g., Perone et al., 1988) that differences in operant performance between humans and other animal species may be due to the fact that different kinds of reinforcement are used and that only "weak" reinforcers are used in human studies. Lowe (1979), however, has presented evidence showing that schedule patterns specific to verbally competent humans are obtained regardless of the type of reinforcer used in different studies. And how do we determine whether a reinforcer is "weak"? Most of the subjects in the present study responded at high rates, approximately 200 to 300 responses per minute, and many frequently expressed an eagerness not just to earn as much money as possible but also to "succeed" and score highly on the task. The fact that this may have led some of them to adopt, for example, an exclusive preference strategy (see Subject CM, Table 1) is clearly not due to a failure of motivation. It should also be noted that in the absence of good empirical support, the "different reinforcer" argument may be a counsel of despair because, as Wearden (1988) has pointed out, it is impossible to "equate" reinforcers between different animal species, let alone between nonhumans and humans.

In our attempts to replicate Bradshaw's findings and also to identify the variables that might favor matching performance in humans, one of the most striking results is that none of our interventions, whether it be a COD, ordinal cue, ordinal information, or some combination of these, produced consistent approximation to matching. Nevertheless, a comparison of the data from Experiments 1, 3, and 5 suggests that gross undermatching or indifference performances were less prevalent when ordinal information about scheduled reinforcement frequencies was provided and that ordinal instructions and cues were also highly correlated with a remarkably rapid onset of matching-like performances (e.g., see Experiments 5 and 6). This is also supported by the schedule-preference data that show that the 11 subjects who ranked the six schedules accurately were in those experiments (3 through 6) that provided ordinal cues; with one exception, none of these had slope values lower than 0.79. Because Bradshaw's studies reliably incorporated ordinal information of this kind, it is possible that this feature may have contributed to the unusually high incidence of matching-like performance that he reports. But even if ordinal cues are implicated in producing greater conformity to matching, this in turn raises the question of how they come to acquire this role. Before this experiment, no one-to-one relation between these particular stimuli and matching behavior had been established in any species, including humans. Despite this it appears that these stimuli occasioned verbal responses pertaining to the concept of ordinality and matching-like behavior in many of the subjects. The verbal community commonly establishes verbal responses of the kind "A is greater (or better) than B is greater than C is greater than . . . Z" in the presence of a variety of spoken or written stimuli such as "1, 2, 3, . . ., 9"; in these cases, the relation between a stimulus and the conventional responses occasioned is usually fairly predictable. On the present evidence, if the six VI schedules comprising Component A had been labeled in order of scheduled reinforcement as 1, 2, 3, 4, 5, and 6, respectively, there is little doubt that the ordinal relation between schedule value and schedule label, together with brief exposure to the scheduled contingencies, would have led many subjects to describe VI 10 s (1) as the richest, VI 20 s (2) as next richest and so on, with VI 720 s (6) as the leanest of the schedules available. In turn, such a description may favor the emergence and maintenance of a matching-like rule for responding on the concurrent schedules. In the case of the ordinal cue stimuli employed in Experiment 3, however, other competing verbal responses appear to have been evoked, which in turn gave rise to different verbal formulations of the contingencies and rules for responding. For example, Subject PB "overinterpreted" the numerical significance of the schedule stimuli; she reported that she used them as a basis for the construction of complex mathematical rules to govern her responding on the operanda. Such ordinal cues are not, therefore, in any simple sense, *textual* verbal stimuli (cf. Skinner, 1957) because they do not reliably occasion a response of a specified form that accords with an arbitrary but conventional relation established by the verbal community.

Because they are not textual or instructional verbal stimuli in the strict sense, consideration of the role of ordinal cues falls outside the scope of much of the literature on human operant behavior (e.g., Baron & Galizio, 1983; Navarick, Bernstein, & Fantino, 1990). For their effect, these stimuli depend upon the human verbal capacity for describing and "analyzing" the experimental situation and subsequently formulating rules for responding; this is a capacity that, dependent as it is on each individual's particular history in the verbal community, is to some degree idiosyncratic (Lowe, 1979).

In addition to the verbal cues that the experimenter either wittingly or unwittingly provides and those that subjects can produce for themselves, there is yet another way in which matching may be generated in a group of subjects. This can occur when 1 subject formulates an accurate description of the contingencies and a corresponding performance rule and communicates these to others in the group. Evidence for this comes from the verbal reports and schedule performance of CL. She reported that she altered her performance rule, and subsequently her schedule behavior, as a result of speaking to another subject (NL, who had a matching performance rule) on the eve of her final session. Indeed, her schedule behavior changed abruptly from indifference ($a = 0.14$) on the penultimate session to matching ($a = 1.12$) on the final day. Another example of this type of verbal control in human choice experiments is reported in Horne (1986). Unless we listen to what subjects like CL have to say to us as experimenters and fully recognize that they *can* say and listen, we will remain ignorant of the "problems" that can arise when subjects talk to each other in human operant experiments. This is not, as far as we are aware, a difficulty encountered in experimentation with nonhumans.

The foregoing account of human choice, relying as it does upon an analysis of the role of verbal behavior and of postexperimental verbal reports, may pose problems for those who object to the inclusion of verbal (including covert) events in a functional analysis. Nonradical behaviorists may argue on theoretical grounds, or radical behaviorists may argue on a pragmatic basis, that even if such events do affect behavior, they cannot be directly observed or changed and thus do not enhance prediction and control (Perone et al., 1988). It may also be argued that the rules articulated in postexperimental questionnaires should not be given causal status; they might not have been operative during the experiment and might have been constructed as post hoc rationalizations of performance (Perone et al., 1988; Shimoff, 1984).

The theoretical justification for including verbal events, both overt and covert, in an analysis of human operant behavior has been admirably set forth by Skinner (1945, 1957, 1974) and will not be repeated here. Skinner has, however, acknowledged that verbal behavior does not readily lend itself to causal analyses because it is itself multiply determined, often covert and, on occasion, incipient in form. Clearly, the analysis of effects of verbal behavior on other human operant performance presents major methodological problems. Postexperimental verbal reports, for example, may be problematic because subjects may not be either able or willing to describe the determinants of their behavior (cf. Nisbett & Wilson, 1977). An interesting feature of the present set of experiments, however, was the close correspondence between the performance rules subjects articulated in their verbal reports and their schedule behavior; this was evidenced not only in the overall correlations but also for each individual. An interesting exception to this was Subject JP in Experiment 2 who, although expressing no preference between the three leanest schedules, nevertheless showed a slight progressive shift in her relative responding from the variable VI component as reinforcement frequency decreased. Much more typical, however, were those subjects whose performance rules and behavior were closely related. Indeed, when such subjects (e.g., PB in Experiment 3) report that they used precise numerical progressions to cue their responding and it is observed that their allocation is consistent with control by these specific sequences, it is difficult to believe that the complex behavioral sequence was generated first and the mathematical formulations came later. A more plausible explanation is that when performing on these concurrent schedules, adult humans will generally attempt to assess the reinforcement schedules in operation and will construct explicit rules for responding that are fairly easily recalled in postexperimental questionnaires. There is no reason to believe, of course, that questionnaires will yield equally valid data in all experiments; whether they do or not may depend as much upon the nature of the questionnaire as the experiment itself. The questionnaire used in this series of experiments (see General Method) is more extensive than those usually employed in human operant experiments and was designed to evoke as much information as possible while minimizing bias of responses.

To argue that post hoc verbal reports in themselves constitute causal variables would, of course, make little sense (cf. Shimoff, 1984); the radical behaviorist view is more subtle and is perhaps best expressed by Skinner himself: "The present analysis . . . continues to deal with the private event, even if only as an inference. It does not substitute the verbal report from which the inference is made for the event itself. The verbal report is a response to the private event and may be used as a source of information about it" (Skinner, 1953, p. 282). This is precisely the approach adopted in the present analysis, which, in order to account for the complexities of human behavior, draws inferences about the existence of rules and covert verbal events from what subjects said and did. In addition to providing a plausible account of the existing data on human choice, an additional merit of the verbal analysis is that it can be subjected to further experimental verification. For example, because humans are sophisticated speaker-listeners, an independent rater may, in accordance with the practices of the verbal community, respond as a listener (or reader) to a subject's contingency descriptions and performance rules and predict the performance that should occur were those rules to be followed. To the extent that such a procedure enhances prediction over and above that based only upon experimenter-provided instructions and scheduled contingencies, the inference of a determining role in operant performance for rules of the kind reported by those subjects is supported (see Zuriff, 1985). Evidence gained in "listening" to subjects' contingency descriptions and performance rules may then be experimentally evaluated by, for example, using their contingency descriptions to refine experimenter-provided instructions or to provide subjects with a particular performance rule that can then be assessed in relation to their operant performance. If subjects' verbalizations are recorded either during or after these interventions, both prediction and control of human operant behavior may be further enhanced and the role of verbal behavior as a determinant of human choice can be systematically evaluated (see Wulfert, Dougher, & Greenway, 1991). Lowe (1983) has outlined a range of experimental strategies designed to test inferences derived from subjects' verbal reports.

Demands that verbal events should not be implicated as "causes" in an experimental analysis unless the causal sequence can be traced to its environmental origins (Hayes, 1986; Perone, 1988; Perone et al., 1988; Shimoff, 1984) can be answered in two ways. First, consider the example of subjects in the present study who used the cues provided in Experiments 3 through 6 to rank ordinally the reinforcement schedules; we have provided an account of how the verbal community establishes the verbal properties of such stimuli. In this general theoretical sense one can, following Skinner (1945, 1957), trace verbal behavior back to early environmental histories within a verbal community. If we require an experimental analysis to give a complete specification of the environmental determinants of the behavior we record, however, this would include in the present case an account of specific environmental interaction in each individual's childhood that led to the development of the concept of ordinality and then the additional experiences with similar cues and equivalent verbal stimuli leading to the eventual rule formulation in the course of our experiment. To expect the experimental analysis of adult human behavior to provide such an account for every empirical undertaking is wholly impracticable.

In summary, the present set of experiments provides evidence for complex interactions among experimental contingencies, verbal behavior, and schedule performance. In attempting to draw attention to the fact that verbal cues and rules have a role to play in studies of this kind, we do not wish to assert that they are the only, or even the most important, determinants of schedule performance; verbal behavior itself clearly has its origins in environmental consequences and is maintained by them. The precise balance between the operation of these and other variables in studies of human choice must be a matter for future research. It may be the case that experimental procedures can be devised that will yield only contingency-shaped choice; this might be achieved, for example, by disguising the contingencies in operation. Clearly, were such an experiment to be conducted, we would expect the findings to conform to the matching law. Demonstrating contingency-shaped behavior in adult humans in laboratory conditions is not, however, an easy task (see Svartdal, 1991). To do so for human choice remains an exciting challenge.

REFERENCES

Baron, A., & Galizio, M. (1983). Instructional control of human operant behavior. *Psychological Record*, **33**, 495–520.

Baron, A., Perone, M., & Galizio, M. (1991). Analyzing the reinforcement process at the human level: Can application and behavioristic interpretation replace laboratory research? *Behavior Analyst*, **14**, 95–105.

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, **22**, 231–242.

Baum, W. M. (1975). Time allocation in human vigilance. *Journal of the Experimental Analysis of Behavior*, **23**, 45–53.

Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior,* **32,** 269–281.

Baum, W. M. (1983). Comment: Matching, statistics, and common sense. *Journal of the Experimental Analysis of Behavior,* **39,** 499–501.

Bem, S. L. (1967). Verbal self-control: The establishment of effective self-instruction. *Journal of Experimental Psychology,* **74,** 485–491.

Bentall, R. P., & Lowe, C. F. (1987). The role of verbal behavior in human learning: 3. Instructional effects in children. *Journal of the Experimental Analysis of Behavior,* **47,** 177–190.

Bentall, R. P., Lowe, C. F., & Beasty, A. (1983). Does language produce behavioural rigidity in humans? *Behaviour Analysis Letters,* **3,** 251–252.

Bentall, R. P., Lowe, C. F., & Beasty, A. (1985). The role of verbal behavior in human learning: 2. Developmental differences. *Journal of the Experimental Analysis of Behavior,* **43,** 165–181.

Bradshaw, C. M. (1977). Suppression of response rates in variable-interval schedules by a concurrent schedule of reinforcement. *British Journal of Psychology,* **68,** 473–480.

Bradshaw, C. M., Ruddle, H. V., & Szabadi, E. (1981). Studies of concurrent performances in humans. In C. M. Bradshaw, E. Szabadi, & C. F. Lowe (Eds.), *Quantification of steady-state operant behaviour* (pp. 79–90). Amsterdam: Elsevier/North Holland Biomedical Press.

Bradshaw, C. M., & Szabadi, E. (1988). Quantitative analysis of human operant behavior. In G. Davey & C. Cullen (Eds.), *Human operant conditioning and behavior modification* (pp. 225–259). New York: Wiley.

Bradshaw, C. M., Szabadi, E., & Bevan, P. (1976). Behavior of humans in variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior,* **26,** 135–141.

Bradshaw, C. M., Szabadi, E., & Bevan, P. (1977). Effect of punishment on human variable-interval performance. *Journal of the Experimental Analysis of Behavior,* **27,** 275–279.

Bradshaw, C. M., Szabadi, E., & Bevan, P. (1979). The effect of punishment on free-operant choice behavior in humans. *Journal of the Experimental Analysis of Behavior,* **31,** 71–81.

Bradshaw, C. M., Szabadi, E., Bevan, P., & Ruddle, H. V. (1979). The effect of signaled reinforcement availability on concurrent performance in humans. *Journal of the Experimental Analysis of Behavior,* **32,** 65–74.

Buskist, W. F., & Miller, H. L., Jr. (1981). Concurrent operant performances in humans: Matching when food is the reinforcer. *Psychological Record,* **31,** 95–100.

Catania, A. C. (1966). Concurrent operants. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 213–270). New York: Appleton-Century-Crofts.

Catania, A. C., Matthews, B. A., & Shimoff, E. (1982). Instructed versus shaped human verbal behavior: Interactions with nonverbal responding. *Journal of the Experimental Analysis of Behavior,* **38,** 233–248.

Catania, A. C., & Reynolds, G. S. (1968). A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior,* **11,** 327–383.

Conger, R., & Killeen, P. (1974). Use of concurrent operants in small group research: A demonstration. *Pacific Sociological Review,* **17,** 399–416.

de Villiers, P. (1977). Choice in concurrent schedules and a quantitative formulation of the law of effect. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 233–287). Englewood Cliffs, NJ: Prentice-Hall.

de Villiers, P. A., & Herrnstein, R. J. (1976). Toward a law of response strength. *Psychological Bulletin,* **83,** 1131–1153.

Dougan, J. D., & McSweeney, F. K. (1985). Variation in Herrnstein's r_0 as a function of alternative reinforcement rate. *Journal of the Experimental Analysis of Behavior,* **43,** 215–223.

Dugdale, N., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 115–138). Hove, England: Erlbaum.

Epling, W. F., & Pierce, W. D. (1983). Applied behavior analysis: New directions from the laboratory. *Behavior Analyst,* **6,** 27–37.

Epling, W. F., & Pierce, W. D. (1988). Applied behavior analysis: New directions from the laboratory. In G. Davey & C. Cullen (Eds.), *Human operant conditioning and behavior modification* (pp. 43–58). New York: Wiley.

Hamblin, R. L. (1979). Behavioral choice and social reinforcement: Step function versus matching. *Social Forces,* **57,** 1141–1156.

Hayes, S. C. (1986). The case of the silent dog—verbal reports and the analysis of rules: A review of Ericsson and Simon's *Protocol Analysis: Verbal Reports As Data. Journal of the Experimental Analysis of Behavior,* **45,** 351–363.

Hayes, S. C., Brownstein, A. J., Haas, J. R., & Greenway, D. E. (1986). Instructions, multiple schedules, and extinction: Distinguishing rule-governed from schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior,* **46,** 137–147.

Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I., & Korn, Z. (1986). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior,* **45,** 237–256.

Hayes, S. C., Zettle, R. D., & Rosenfarb, I. (1989). Rule-following. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 191–220). New York: Plenum Press.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior,* **13,** 243–266.

Herrnstein, R. J. (1974). Formal properties of the matching law. *Journal of the Experimental Analysis of Behavior,* **21,** 159–164.

Horne, P. J. (1986). *Human choice and the matching law: The role of verbal behaviour.* Unpublished doctoral dissertation, University College of North Wales, Bangor, Wales, UK.

Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology,* **1,** 134–148.

LeFrancois, J. R., Chase, P. N., & Joyce, J. H. (1988). The effects of variety of instructions on human fixed-interval performance. *Journal of the Experimental Analysis of Behavior,* **49,** 383–393.

Levine, M. (1971). Hypothesis theory and nonlearning despite ideal S-R reinforcement contingencies. *Psychological Review,* **78,** 130–140.

Lewis, D. (1960). *Quantitative methods in psychology.* New York: McGraw-Hill.

Lowe, C. F. (1979). Determinants of human operant behavior. In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behaviour: Vol. 1. Reinforcement and the organization of behavior* (pp. 159-192). Chichester, England: Wiley.

Lowe, C. F. (1983). Radical behaviorism and human psychology. In G. C. L. Davey (Ed.), *Animal models of human behavior: Conceptual, evolutionary, and neurobiological perpectives* (pp. 71-93). Chichester, England: Wiley.

Lowe, C. F., Beasty, A., & Bentall, R. P. (1983). The role of verbal behavior in human learning: Infant performance on fixed-interval schedules. *Journal of the Experimental Analysis of Behavior,* **39,** 157-164.

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behaviour.* New York: Liveright.

Matthews, B. A., Shimoff, E., Catania, A. C., & Sagvolden, T. (1977). Uninstructed human responding: Sensitivity to ratio and interval contingencies. *Journal of the Experimental Analysis of Behavior,* **27,** 453-467.

McDowell, J. J. (1981). On the validity and utility of Herrnstein's hyperbola in applied behaviour analysis. In C. M. Bradshaw, E. Szabadi, & C. F. Lowe (Eds.), *Quantification of steady-state operant behaviour* (pp. 311-324). Amsterdam: Elsevier/North Holland [Biomedical Press].

McDowell, J. J. (1982). The importance of Herrnstein's mathematical statement of the law of effect for behavior therapy. *American Psychologist,* **37,** 771-779.

McSweeney, F. K., Melville, C. L., Buck, M. A., & Whipple, J. E. (1983). Local rates of responding and reinforcement during concurrent schedules. *Journal of the Experimental Analysis of Behavior,* **40,** 79-98.

Michael, R. L., & Bernstein, D. J. (1991). Transient effects of acquisition history on generalization in a matching-to-sample task. *Journal of the Experimental Analysis of Behavior,* **56,** 155-166.

Navarick, D. J., Bernstein, D. J., & Fantino, E. (1990). The experimental analysis of human behavior. *Journal of the Experimental Analysis of Behavior,* **54,** 159-162.

Navarick, D. J., & Chellsen, J. (1983). Matching versus undermatching in the choice behavior of humans. *Behaviour Analysis Letters,* **3,** 325-335.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review,* **84,** 231-259.

Oscar-Berman, M., Heyman, G. M., Bonner, R. T., & Ryder, J. (1980). Human neuropsychology: Some differences between Korsakoff and normal operant performance. *Psychological Research,* **41,** 235-247.

Paniagua, F. A., & Baer, D. M. (1982). The analysis of correspondence training as a chain reinforceable at any point. *Child Development,* **53,** 786-798.

Perone, M. (1988). Laboratory lore and research practices in the experimental analysis of human behavior: Use and abuse of subjects' verbal reports. *Behavior Analyst,* **11,** 71-75.

Perone, M., Galizio, M., & Baron, A. (1988). The relevance of animal-based principles in the laboratory study of human operant conditioning. In G. Davey & C. Cullen (Eds.), *Human operant conditioning and behavior modification* (pp. 59-85). Chichester, England: Wiley.

Pierce, W. D., & Epling, W. F. (1980). What happened to analysis in applied behavior analysis? *Behavior Analyst,* **3,** 1-9.

Pierce, W. D., & Epling, W. F. (1983). Choice, matching, and human behavior: A review of the literature. *Behavior Analyst,* **6,** 57-76.

Pierce, W. D., Epling, W. F., & Greer, S. M. (1981). Human communication and the matching law. In C. M. Bradshaw, E. Szabadi, & C. F. Lowe (Eds.), *Quantification of steady-state operant behaviour* (pp. 345-348). Amsterdam: Elsevier/North Holland Biomedical Press.

Poppen, R. (1982). Human fixed-interval performance with currently programmed schedules: A parametric analysis. *Journal of the Experimental Analysis of Behavior,* **37,** 251-266.

Rachlin, H. (1980). Economics and behavioral psychology. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 205-236). New York: Academic Press.

Ruddle, H., Bradshaw, C. M., Szabadi, E., & Bevan, P. (1979). Behaviour of humans in concurrent schedules programmed on spatially separated operanda. *Quarterly Journal of Experimental Psychology,* **31,** 509-517.

Schmitt, D. R. (1974). Effects of reinforcement rate and reinforcer magnitude on choice behavior of humans. *Journal of the Experimental Analysis of Behavior,* **21,** 409-419.

Schroeder, S. R. (1975). Perseveration in concurrent performances by the developmentally retarded. *Psychological Record,* **25,** 51-64.

Schroeder, S. R., & Holland, J. G. (1969). Reinforcement of eye movement with concurrent schedules. *Journal of the Experimental Analysis of Behavior,* **12,** 897-903.

Schwartz, B. (1984). *Psychology of learning and behavior* (2nd ed.). New York: W. W. Norton.

Shimoff, E. (1984). Post-session questionnaires. *Experimental Analysis of Human Behavior Bulletin,* **2,** 1.

Shimoff, E., Matthews, B. A., & Catania, A. C. (1986). Human operant performance: Sensitivity and pseudosensitivity to contingencies. *Journal of the Experimental Analysis of Behavior,* **46,** 149-157.

Siegel, S., & Castellan, N. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Silberberg, A., Thomas, J. R., & Berendzen, N. (1991). Human choice on concurrent variable-interval variable-ratio schedules. *Journal of the Experimental Analysis of Behavior,* **56,** 575-584.

Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review,* **52,** 270-277.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Skinner, B. F. (1957). *Verbal behavior.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1974). *About behaviorism.* New York: Knopf.

Staddon, J. E. R. (1968). Spaced responding and choice: A preliminary analysis. *Journal of the Experimental Analysis of Behavior,* **11,** 669-682.

Svartdal, F. (1991). Operant modulation of low-level attributes of rule-governed behavior by nonverbal contingencies. *Learning and Motivation,* **22,** 406-420.

Takahashi, M., & Iwamoto, T. (1986). Human concurrent performances: The effects of experience, instructions, and schedule-correlated stimuli. *Journal of the Experimental Analysis of Behavior,* **45,** 257-267.

Taylor, R., & Davison, M. (1983). Sensitivity to reinforcement in concurrent arithmetic and exponential schedules. *Journal of the Experimental Analysis of Behavior,* **39,** 191–198.

Timberlake, W. (1982). The emperor's clothes: Assumptions of the matching theory. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts* (pp. 549–568). Cambridge, MA: Ballinger.

Todorov, J. C., Oliveira Castro, J. M., Hanna, E. S., Bittencourt de Sa, M. C. N., & Barreto, M. Q. (1983). Choice, experience, and the generalized matching law. *Journal of the Experimental Analysis of Behavior,* **40,** 99–111.

Warren-Boulton, F. R., Silberberg, A., Gray, M., & Ollom, R. (1985). Reanalysis of the equation for simple action. *Journal of the Experimental Analysis of Behavior,* **43,** 265–277.

Wearden, J. H. (1988). Some neglected problems in the analysis of human operant behavior. In G. Davey & C. Cullen (Eds.), *Human operant conditioning and behavior modification* (pp. 197–224). New York: Wiley.

Wearden, J. H., & Burgess, I. S. (1982). Matching since Baum (1979). *Journal of the Experimental Analysis of Behavior,* **38,** 339–348.

Weiner, H. (1965). Conditioning history and maladaptive human operant behavior. *Psychological Reports,* **17,** 935–942.

Weiner, H. (1969). Controlling human fixed-interval performance. *Journal of the Experimental Analysis of Behavior,* **12,** 349–373.

Wilkinson, G. N. (1961). Statistical estimations in enzyme kinetics. *Biochemical Journal,* **80,** 324–332.

Winkler, R. C. (1980). Behavioral economies, token economies and applied behavior analysis. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 269–297). New York: Academic Press.

Wulfert, E., Dougher, M. J., & Greenway, D. E. (1991). Protocol analysis of the correspondence of verbal behavior and equivalence class formation. *Journal of the Experimental Analysis of Behavior,* **56,** 489–504.

Wurster, R. M., & Griffiths, R. R. (1979). Human concurrent performances: Variation of reinforcer magnitude and rate of reinforcement. *Psychological Record,* **29,** 341–354.

Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction.* New York: Columbia University Press.

Received January 17, 1992
Final acceptance July 31, 1992